SOLDIER SELECTION

Past, Present, and Future

Lola M. Zook

1996

DTIC QUALITY INSPECTED 3

**United States Army Research Institute
for the Behavioral and Social Sciences**

19970228 057

DISCLAIMER NOTICE

THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE
COPY FURNISHED TO DTIC
CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO
NOT REPRODUCE LEGIBLY.

SOLDIER SELECTION
Past, Present, and Future

Lola M. Zook

1996

DTIC QUALITY INSPECTED 2

**United States Army Research Institute
for the Behavioral and Social Sciences**

REPORT DOCUMENTATION PAGE		Form Approved OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE February 1996	3. REPORT TYPE AND DATES COVERED Final Report 20 Sep 94 – 14 Feb 96

4. TITLE AND SUBTITLE

Soldier Selection: Past, Present, and Future

5. FUNDING NUMBERS

MDA903-93-D-0032, DO 25
62785
A790
1211
C06

6. AUTHOR(S)

Lola M. Zook

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Human Resources Research Organization
66 Canal Center Plaza
Suite 400
Alexandria, VA 22314

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Institute for the Behavioral and
   Social Sciences
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

Special Report 28

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This Special Report is intended for readers who have an interest in military personnel selection and assignment. It describes a long-term research program—Project A and Building the Career Force—to evaluate and enhance the Army's process of selecting qualified applicants for enlistment and assigning them to the most appropriate jobs. A model of the elements of enlisted job performance is presented. Then findings on the powers and limits of selection and assignment tests, both operational and experimental, are reported. This research shows that it is hard to improve on the power of the operational test battery (the Armed Services Vocational Aptitude Battery [ASVAB]) to predict individuals' overall skilled technical performance in the Army, but that experimental tests of temperament add greatly to predicting motivational outcomes such as effort, discipline, and completion of enlistment. Applicants' scores on ASVAB also relate strongly to their performance in combat and in the second tour. An extensive reference list and glossary of terms in selection and assignment are provided. Appendices depict the enlisted career life cycle and (cont. on back)

14. SUBJECT TERMS Personnel selection; Personnel classification Job performance; Performance modeling; Armed Services Vocational Aptitude Battery; Project A; (cont. on back)	15. NUMBER OF PAGES 102
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

13. ABSTRACT Cont.
a give short history of military selection.

14. SUBJECT TERMS Cont.
Building the Career Force; Validation; Predictor development

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

Contents

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

Contents (Continued)

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

List of Figures and Tables

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

FOREWORD

Before beginning their Army enlisted careers, civilian volunteers deal with a system for selecting the best applicants and matching them with Army jobs. This volume describes a long-term program of research to evaluate the existing system and to develop and try out possible enhancements to it. The goal is to enable the highest level of Army performance by maintaining the highest quality in personnel.

At the core of the system is the Armed Services Vocational Aptitude Battery (ASVAB), which measures verbal, quantitative, and technical aptitudes. Our confidence in ASVAB today rests on the long history of research and development behind it. The present report summarizes that history, confirms the effectiveness of ASVAB, and describes the promise of new measures of varied human attributes.

Because most applicants do not have much of a job history to evaluate, their scores on ASVAB serve as surrogate measures, or **predictors**, of their future performance in the Army. The enhancements to ASVAB that were evaluated here were designed to bring us closer to measuring the whole person. They include measures of both aptitudes (psychomotor and spatial) and dispositions (temperament and vocational interests).

In this special report you will find that both the nature of Army job performance and the personal qualities that underlie that performance are complex. The very concept of *quality in Army personnel* is a tool of manpower management that requires, as the Army of the future evolves, maintenance and renewal. This report describes a very ambitious research program that is proving to be a major part of the renewal of the conception of personnel quality and of the Army's steps to make the best use of that quality.

Edgar M. Johnson
Director
U.S. Army Research Institute for the
Behavioral and Social Sciences

ACKNOWLEDGMENTS

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

INTRODUCTION

This special report is about *selection* and *classification*, or picking the best persons for the enlisted Army and then matching them with the Army jobs that require their individual talents. Here we present the findings of a long-term research and development program on the subject at the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI). Because this program involved the participation of over 60,000 soldiers[1], won many awards, and is having an impact on testing for all of the Services, its name -- Project A (aimed at the first enlistment) and Building the Career Force (aimed at the second) -- may be familiar to you. Called Project A/Career Force for short, this program had two major objectives:

- To evaluate the effectiveness of the Armed Services Vocational Aptitude Battery (ASVAB) in distinguishing applicants who will go on to perform well in Army jobs from those who will not; and

- To identify human abilities and dispositions that are not measured by ASVAB but that are likely to underlie success in Army jobs, develop tests of those attributes, and then evaluate the tests' power to predict examinee's' relative success in actual Army jobs. The yardstick for judging the potential benefit of the new tests was the extent to which they could add to ASVAB in predicting soldiers' job performance.

A point-by-point summary of this program's findings is given on pages xi and xii. To put the findings in context, Chapter 1 of this report describes the sequence of events as individuals move from being recruited for initial enlistment all the way up to re-enlistment. In Chapter 2, we relate the objectives of this research to the sometimes competing interests of stakeholders in the operational and management sides of the Army.

To evaluate how well a personnel test predicts examinee's' later performance in Army jobs, one must start with a conception of that performance and an adequate way to measure it. Chapter 3 presents our research answer to the fundamental question "What does Army job performance consist of?" In Chapter 4 we describe the new Project A/Career Force tests, then summarize how good we found ASVAB and the new tests to be for predicting first-tour performance. Chapter 5 elaborates the findings, noting that the "can do" and "will do" components of performance rely on different human attributes. Also, Chapter 5 extends the time horizon for prediction to the second tour, where leadership emerges as a major aspect of performance. In Chapter 6 we present our findings on attrition, retention, and performance in combat. Chapter 7 discusses issues in realizing the potential benefits of personnel tests. In closing, Chapter 8 looks to the future.

Readers who wish to consult technical sources on Project A/Career Force will find an extensive bibliography after the text. Following the reference list, key technical terms in selection, classification, and military personnel testing are defined in a Glossary. Supplementing the treatment of the personnel life cycle in Chapter 1, Appendix A depicts that cycle in a more detailed flow-chart form. Appendix B gives a short history of military selection.

---

[1] As examinees, test administrators, subject matter experts, sponsors, and scientific and policy reviewers.

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

The research program described here was the largest validation of a personnel test that has ever been conducted. It confirmed that the Army's present selection system is effective, contributed enhancements to the existing system, and helped point the way to the future with new experimental tests. Based on the rich experience of this program and related ARI research, we are can confidently say "Here's what we know about selection and classification."

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

Preview: Summary of Findings About Soldier Selection[1]

About Soldier Performance

- First-tour enlisted performance is multi-dimensional, consisting of the following elements:

  - Core Technical Proficiency
  - General Soldiering Proficiency
  - Effort and Leadership
  - Personal Discipline
  - Physical Fitness and Military Bearing

- In order to know how well we have selected the right persons and matched them with the right jobs, we must be able to measure these elements.

- The two proficiency dimensions--Core Technical and General Soldiering--are the "Can-Do" element of performance, showing that the soldier is able to perform these duties.

- The other three dimensions--Effort and Leadership, Personal Discipline, and Physical Fitness--are the "Will-Do" element, showing that the soldier is willing to perform effectively over time.

About ASVAB

- Pre-enlistment ASVAB strongly predicts "Can-Do" performance -- proficiency in MOS-specific and common tasks and duties.

- ASVAB does not do as well in predicting Personal Discipline, Fitness, and Military Bearing, where motivation has more impact.

- The ASVAB has very good coverage of verbal and math aptitudes, but does not measure such other attributes as temperament or eye-hand coordination.

About Prediction and Performance

- Each element of performance relies on somewhat different human capabilities and dispositions.

- There is no single best enlistment test battery. The combination of subtests that is best for one purpose, such as classification efficiency, often is not the best for other purposes. Tradeoffs must be made, depending on the specific purposes.

- The ASVAB is the best predictor of proficiency on the job. Its Quantitative and Technical components add the most to predicting General Soldiering Proficiency, and the Verbal component is the best for predicting Core Technical Proficiency.

- Both spatial ability and the perceptual-psychomotor tests predict job proficiency well. However, if

added to the ASVAB, they produce just a small gain over ASVAB alone.

- The attribute of Dependability (based on such ABLE scales as Conscientiousness and Nondelinquency) has the highest relationship with Personal Discipline, a "Will-Do" dimension of performance.

- ASVAB contributes the most to predicting the Effort and Leadership dimension for first-tour soldiers.

- Performance in combat (Operation Desert Shield/Storm) was highly related to peacetime performance, both before and after the conflict. The ASVAB, at enlistment, predicted task proficiency and supervision during combat.

- Comparisons of performance by soldiers at the several aptitude levels show that higher aptitude soldiers perform better in general on a variety of performance indicators.

About Attrition and Second Tour

- The relationship between attrition and measures of cognitive ability, such as AFQT and other ASVAB composites, is not very strong.

- Addition of ABLE temperament scores, such as the Nondelinquency scale, add substantially to ASVAB and education in identifying the attrition-prone.

- The elements of performance remain much the same as the soldier gains experience. Leadership emerges as an additional element in the second tour.

- Performance in training predicts performance in first-tour jobs, and first-tour performance predicts performance of second-tour NCOs.

- Pre-enlistment ASVAB scores are strongly related to second-tour performance, including Leadership.

- The new tests from Project A also add unique prediction of second-tour performance. Thus, leadership is based on a combination of personal qualities.

---

[1]See the Glossary for explanations of terms. ASVAB is the Armed Services Vocational Aptitude Battery. ABLE is a self-report test measuring traits of temperament.

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

Who is Selected? and Who Succeeds?

The Ultimate Smart Weapon: The U.S. Soldier

The first and most important element in an army is the soldier. And this next-door neighbor kind of person is also the element that has the greatest potential for success or for failure as the Army goes about its work. No matter how advanced the weaponry, no matter how inspired the strategy, success lies with the skills and determination of all individual soldiers out to the end of the chain of command.

Fully aware of the overriding value of the human element in the system, the U.S. Army has long invested heavily -- in both time and money -- in producing top-level soldiers.

Pursuit of this goal involves four main approaches: selection, classification, training, and leadership. Their nature and emphasis change over time, reflecting changes in Army missions and in our society as a whole. Currently, these influences range from the end of the Cold War to the shifting mosaic of the general population from which the Army draws its members. Demands on the Army to change with the times have seldom been as complex.

Much of what is written in and about the Army focuses on training and leadership. Less often do we read about selecting individuals for entry into the Army or assigning them to any one of over 200 entry-level jobs -- that is, about *selection* and *classification*. Yet these steps control everything that follows.

Recruits are usually regarded as a given, the raw material to be turned into competent soldiers. But the nature of the raw material is far from being a given. It varies greatly depending on the civilian job market for youth, the Army's requirements for aptitude and education in new recruits, and the latest technologies for selecting and assigning on the basis of each individual's abilities. Such

---

An Army raised without proper regard to the choice of its recruits was never made good by length of time.
Vegetius, *De Re Militari*, A.D. 378

We simply cannot do everything we have to do without quality soldiers.
GEN Gordon R. Sullivan
Chief of Staff, United States Army

---

intangibles will influence who enters the Army, and who succeeds or fails at the hands of those who train and, ultimately, of those who command.

So, what is being done to learn about young persons who show an interest in the Army? Some youths seem obviously to be well qualified, but can we pick out which ones are likely to fit into Army life? Can we recognize the signs that say this high school senior would fail if assigned to one Army job field but be a solid success in another? Can we look at this youngster in jeans and see the makings of a military leader?

By answering such questions, selection and classification contribute to the later success of training and leadership. The Army and the

other Services have done much research over the years to provide the basis for selecting the best applicants for military service. The purpose of this report is to describe some of what has been learned, or is still being studied, in such research. The objective is to enable both military and civilian readers to understand more fully what it means to be "a good soldier."

The remainder of this chapter summarizes the changing military situations that in turn have created changing requirements for personnel research. It also sketches the Army personnel life cycle, starting with recruiting. A flow diagram of that life cycle is given in Appendix A to this report. The main areas of the present research inquiry are outlined in Chapter 2, along with a look at research concepts as applied in a military setting.

Some of what we now know about soldier selection -- about what enters into "good performance," about aptitudes, about predicting individual success or failure, about attrition, and more -- is presented in Chapters 3-6. Those sections summarize the findings of a wide variety of projects. The thread is continued in Chapter 7, where special attention is paid to potential changes, prospective benefits versus costs, and further opportunities for meaningful inquiry. Chapter 8 looks to the future.

The Military Context for Selection Research

In choosing among youth who want to join the Army, it helps to know that:

- Applicants' scores on the standard enlistment battery, which tests verbal, math, and other aptitudes, are strongly related to later proficiency in Army

tasks -- that is, the tests "predict" effective performance on the job.

- Scores on the current enlistment tests do not do as well in predicting attrition (failure to finish the first enlistment).

- Addition of other types of tests, measuring such things as eye-hand coordination or dependability, can provide useful information -- for example, whether an individual is likely to stay in the Army, or which types of jobs suit this individual.

- "Performance" is not a single thing but has distinct dimensions (e.g., technical proficiency, personal discipline) that are related to success in different job activities and that can be measured in a variety of ways.

These are a few examples of things we now know from research -- things that we did not know in 1981 (though guesses were plentiful). In that year, in an effort to firm up a basis for manpower planning, the Congress directed the Armed Services to "develop a better database on the relationship between factors such as high school graduation and entrance test scores, and effective performance.[46] The research that followed on this directive yielded the findings cited previously and much more, as will be described in later chapters.

At the time of the directive, the Army, which was completing its transition to an all-volunteer force in the drawdown after the Vietnam war, was already expanding research in personnel matters. Times were indeed difficult. The Army was falling short of its manpower goals and depending more on applicants with low aptitudes. Only one-fourth of the 1980 recruits were from the upper half of youth in aptitude, and barely one-half had graduated from high school. There were other formidable problems:

- Technological demands on soldiers were multiplying as the Army launched the largest equipment modernization program since World War II. Anti-tank weapons were becoming wire-guided missiles; tanks were adding onboard computers for gunnery and navigation -- technical change was everywhere.

- The youth population was both shrinking and changing. A decline in numbers of up to 20 percent was projected to stretch on through the mid-1990s. And a larger part of the pool would be youths having English as a second language and/or less education than in the past.

- Critics were questioning the fairness of tests in general, charging that test content often showed cultural, racial, or gender bias. Enlistment tests were not exempt from this barrage.

- And, in 1980, it was discovered that the versions of the Department of Defense test battery that had been used since 1976 to determine eligibility to enlist had been misnormed, so that scores in the lower ranges were artificially inflated. As a result, half of Army recruits in that period were drawn from the bottom 30 percent of the eligible youth pool.[46] All told, 300,000 young men and women who did not really meet enlistment standards entered military service between 1976 and 1980.

Fifteen years later, conditions are both different and similar, depending on where you look. The supply of aptitude/education has vastly improved. As of 1995, 70 percent of Army recruits are from the top half of applicants in aptitude and 95 percent are high school graduates.[49] But the competition for youth of high quality and the Army's need for recruits with strong technical capabilities are stronger than ever.

The change that is perhaps the most difficult to define is in the Army's missions. The end of the Cold War has lessened the need for large-scale land forces and resulted in substantial declines in the size of the Army and the military budget. Yet this situation is paired with greater political instability and military conflict all over the world -- think of Yugoslavia, Somalia, Iraq, and Haiti. Increasingly, peacekeeping duty has taken more of our soldiers' time. To quote General Gordon R. Sullivan, Army Chief of Staff, in 1993:

> Even though the Cold War ended, peace has not broken out all over. There are [fewer] soldiers and less dollars, but plenty of missions. Our pace of operations is running higher than it did during most of the Cold War. We have some 20,000 to 25,000 soldiers overseas on contingencies in 70 to 75 countries, and that's in addition to the 125,000 forward based in Europe, Asia, and elsewhere. . . . We have more soldiers getting shot at right now than at any time during the Cold War except the hottest stretches of the Korean, Vietnam, and Gulf Wars. That's the reality of today's Army.

At the same time, the Army's long-standing mission of humanitarian assistance has been highly visible in recent years -- at home following the Los Angeles riots, Hurricane Andrew in Florida, and two earthquakes in California, for example, as well as overseas. The capability to meet the needs for operations other than war must be maintained in spite of manpower and monetary cuts. In fact, it will require even greater flexibility and ingenuity of the troops placed in such situations.

Other problems remain, although perhaps with a new twist. Political and economic vagaries pose multiple uncertainties for Army planners. Even with lower manpower goals, recruiting is anything but easy in view of the complexity of changing Army requirements. More than ever, the Army needs to find the right people and place them in the right jobs -- and it must update what "right" means as the world changes.

What Are We Doing Now to Select?

During much of the 1980s, which were marked by increased concern over how to select the volunteers who would become top-quality soldiers, the strength of the Army enlisted force was about 800,000. Every year, 300,000 to 400,000 possible recruits were screened, and every year 120,000 to 140,000 young men and women enlisted. The numbers are smaller now, but the process of Army renewal and the problems of trying to choose "the best" are much the same -- and perhaps even more important.

The Life Cycle in the Army

For a civilian applicant, selection is the gateway to becoming a soldier, but it is also much more. Selection continues operating throughout an Army career, always in the interests of getting and keeping "the best." It underlies the assignment of recruits to particular jobs; it is implicit in the pace and even the fact of promotions, one by one over the years; it determines whether a first-termer is allowed to reenlist, and how long the career soldier will serve. The phases in the Army life of the career soldier are summarized in this section and illustrated in Figure A.1. The early stages--recruitment, selection, classification, and assignment--are interwoven and overlapping, and are discussed more fully in the next section.

Selection and Classification. The Army draws primarily upon 17- to 21-year-old youth to fill its personnel needs. From the Army's perspective, the most desirable group is graduating high school seniors and recent graduates.[44] As much as possible, the Army recruits from the higher aptitude members of that group.

Prospects are prescreened by recruiters on such major eligibility points as age and education, and usually are given a short aptitude-type test. Promising applicants are sent to one of the sites operated by the Department of Defense to test applicants for all the Armed Services. Army applicants who pass the aptitude, physical, and other (e.g., moral) screening requirements then explore the various options that can lead to Army enlistment and job assignment.

The great majority of enlistees enter the Army under a specific option that guarantees field assignment, or unit assignment. Upon enlistment, the new recruit will be placed in the Delayed Entry Program (DEP) until the date of the promised training slot.

For the enlistees, this decision will determine the nature of their initial training and occupational assignment, their future military work environment, and their chances of advancement in an Army career. For the Army, the effectiveness of the selection and classification process will significantly affect the level of performance and attrition for the entire force.[11]

Training. After brief processing at an Army Reception Center, recruits enter a Basic Training (BT) program of eight weeks, then Advanced Individual Training (AIT) to provide entry-level MOS skills. Entrants into some of the combat arms and the military

A Unique Requirement

Effective personnel selection and classification is probably more important for the Military Services than for any other known organization, public or private. This is because a principal goal of the Military Services is to maintain a viable defense organization consisting of hundreds of thousands of people who are well qualified, well trained, and expert. Should a critical need arise, the organization will be ready to meet it no matter how unexpected. . . In this context, effective personnel selection and classification is much more difficult in the military than in other public or private organizations. Most individuals who enter have little prior training or work history. Most applicants know almost nothing about the kinds of jobs that are available . . . . Nevertheless, during a relatively short space of time the individual and the organization must evaluate each other and make an occupational choice. Further, for enlisted personnel, the choices must be such that the exact number of available training slots for each job must be filled on a specific date (the day the class starts). All this must take place within a set of complex constraints imposed by the organization and the Congress, not the least of which is spending exactly the amount budgeted for a given period.

Building a Joint-Service Classification Roadmap[41]

police train at one Army post in courses that last three to four months; entrants in other specialties attend separate Army technical schools with course lengths depending on the technical complexity of the MOS. The course diversity is evident from the fact that, as of 1995, the Army has over 200 entry-level MOS.

Most enlisted trainees no longer receive school grades at the end of the course but are evaluated on a Pass/Fail basis. Those who fail certain portions of the course are recycled; the idea is that slower learners, given enough time and effort under self-paced programs, can normally be trained to satisfactory levels. Those who continue to fail may be assigned to another, less technical specialty or discharged.[11]

The First Term of Service. Following AIT, most recruits are assigned to an Army unit for on-the-job or unit training. Additional courses may be given, and there are opportunities to volunteer for special programs (an example of the standards for entering such a program, Ranger training, is included in Figure A.1). Evaluation of a soldier's performance, both on and off the job, is an ever-present part of the picture. Its effects are to be seen in contrasting directions -- attrition and promotions.

Attrition is costly to the Army. Once the soldier is partly trained, the earlier the attrition occurs the greater the loss, since the Army will have paid all the "up-front" expenses of recruiting, accession, and early training without receiving a payback period of service from a trained soldier. Some separations, of course, are nobody's fault and could not have been predicted--family deaths or crises, medical disabilities, and the like. Losses that are directly related to the behavior or character of the person -- serious disciplinary infractions, poor performance, inability to adapt to the military way of life -- are another matter. Improving screening and testing so that more of these unsuitable individuals will be kept from enlisting is a major concern for Army personnel planners.

A soldier's promotion rate is tangible evidence of success. During an initial 3-year enlistment term, the typical enlistee can expect to reach pay grade E-4, although further advancement is possible.

Yes?

No?
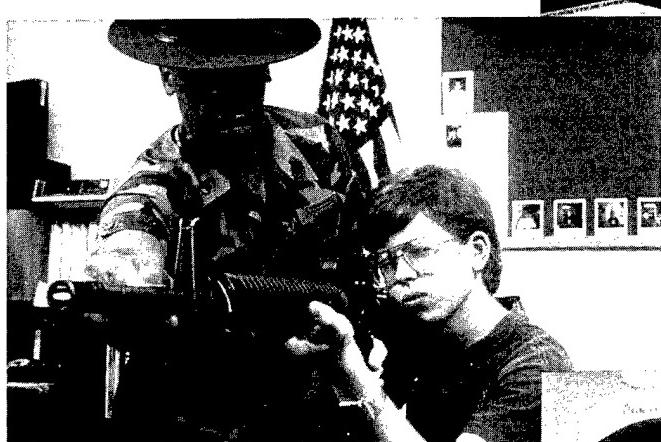
Maybe?

The Second Term of Service. Keeping highly qualified soldiers in the Army beyond their first tour ranks high among Army priorities.[6] A substantial investment has been made in their training, and their skills and experience contribute heavily to the Army's technical competence. Even more important, they join the pool of sergeants who train and lead younger and newer soldiers. On the other hand, soldiers who have performed at a marginal level will not qualify to reenlist.

At any particular time, many indirect factors influence reenlistment. The need for noncoms may vary with technological gains or obsolescence in an MOS, or changes in Army missions. And, in turn, economic conditions in the outside world may make an eligible soldier either more or less interested in reenlisting. The cumulative losses of attrition, reenlistment screening, and non-reenlistment of eligible personnel generally lessen the initial Army cohort to 10 to 20 percent of their original numbers by the time they enter their sixth year of service.[11] To a great extent, however, the Army has met its goal of keeping "the best."

Because the military functions entirely as a "promote from within" organization, it is important that new personnel (enlisted and officer) continuously feed into the system. To this end, the system is engineered to ensure a continuous turnover (non-reenlistment, retirement). It is necessary and expected that a large proportion of first-termers will not continue into a second term. The key is to ensure that those who do reenlist are at high levels of both technical and leadership skills and potential.

Getting Into the Army

Now for a closer look at the part of the Army life cycle--selection and classification--

that is our primary topic. The first flicker of action in the selection process occurs when a young person (often a high school senior) walks into an Army recruiting office. This is not, however, where the Army effort starts. The Army has been reaching out through TV, newspaper, and magazine ads, school Career Days, and high school ROTC units. All this effort has been aimed at interesting members of the high-quality youth market, in competition with the other Services, businesses, and colleges. A useful assessment tool for schools and students, as well as for the Services, is the paper-and-pencil form of the ASVAB that is given to 900,000 high school juniors and seniors annually in the DoD Student Testing Program.

Early Exploration. The recruiter tells the prospect about the enlistment options to Army service, at the same time doing a "prescreening" to sort out those who do not qualify (e.g., on age) and to find out whether the prospect has (or will have) a high school diploma (the Army requires higher test scores for prospects who lack this credential). Prospects who did not take the ASVAB in high school are usually given one of two short tests -- the Computerized Adaptive Screening Test (CAST) or the Enlistment Screening Test (EST) -- which give a good indication of how well the candidate would score on the full ASVAB. Recruiters can then avoid spending their resources on prospects who are unlikely to qualify for service; also, there are limits on the percentage of non-graduates that can be recruited.

If the prescreening looks promising, the more formal steps of the enlistment process follow (see Appendix A). Since the Army pays all applicant costs associated with enlistment processing, such as travel, room, and board,

costs are reduced by prescreening applicants with the CAST or EST.[44]

The ASVAB and the AFQT. The ASVAB -- Armed Services Vocational Aptitude Battery -- is a collection of ten cognitive subtests measuring verbal, math, and technical aptitudes (see Table 1.1). All the Services have used it since 1976 (with modifications) as their primary selection and classification tool. It is the successor of the varied tests used to screen and assign military applicants since World War I days.

The Army's basic requirement for applicant quality is in terms of the Armed Forces Qualification Test (AFQT), a subset of four verbal and math ASVAB subtests. The scores are grouped in percentiles representing the range of scores in the national youth population from which the Army draws its members (see Table 1.2).

AFQT Categories I and II signify well-above and above average trainability, respectively. Category III is the average range, and IV is below average. Individuals scoring in Category V are, by law, not eligible to enlist. The Army seeks to recruit as many as possible from Categories I-IIIA because they are likely to succeed academically in training.[30] The percentage of recruits in Category IV has been 10 percent or less since 1984, 5 percent or less since 1987, and down to 1 percent in 1992. At the same time the Army increased the proportion of accessions in the upper half of the AFQT; in 1992, 75 percent of those accepted were in Category IIIA or higher.[35]

ASVAB subtests are grouped not only into the AFQT composite, but also into ten Aptitude Area Composites, sets that have proved to be best at predicting success in

Table 1.1
Subtests in the Armed Services <u>Vocational Aptitude Battery (ASVAB)</u>

Subtest	Number of Items
General Science (GS)	25
* Arithmetic Reasoning (AR)	30
* Paragraph Comprehension (PC)	15
* Word Knowledge (WK)	35
Numerical Operations (NO)	50
Coding Speed (CS)	84
Auto & Shop Information (AS)	25
* Mathematical Knowledge (MK)	25
Mechanical Comprehension (MC)	25
Electronics Information (EI)	20

* Included in the Armed Forces Qualification Test (AFQT).

Table 1.2
<u>Armed Forces Qualification Test (AFQT)</u>

AFQT Category	Percentile Score Range Population	Percent of Civilian Youth
I	93-99	8
II	65-92	28
IIIA	50-64]	
IIIB	31-49]	34
IVA	21-30]	
IVB	16-20]	21
IVC	10-15]	
V	1-9	9

specific Army training programs. For example, the Aptitude Area score for signal maintenance specialties is summed from the scores on the Arithmetic Reasoning, Mathematical Knowledge, Electronic Information, and General Science subtests.

Figure 1.1. From Civilian to Soldier

AFQT Prescreening (CAST) (Recruiting Station) → Pass Screen — Yes → Took ASVAB (METS/MEPS) — Yes 235,000 → AFQT Qualified — Yes 201,000 →

Pass Screen — No → Screened Out

Took ASVAB (METS/MEPS) — No → Not Tested

AFQT Qualified — No → AFQT Unqualified 34,000

→ Took Physical (MEPS) — Yes 129,000 → Physically Qualified? — Yes 107,000 → Classified Into MOS — Yes →

Took Physical (MEPS) — No → AFQT Qualified Did Not Take Physical 72,000

Physically Qualified? — No → Physically Unqualified 22,000

Classified Into MOS — No → Qualified Did Not Enlist 24,000

→ Enter DEP 83,000 → Enter Duty — Yes 77,000 → Assigned to MOS Training / Assigned to CMF Training → Assigned to MOS

Note: These numbers for a notional recent year show the extent of screening at each step in the beginning of the personnel life cycle. These are interpolated and estimated from known data for FY 93 on signing contracts (i.e., entering the Delayed Entry Program), and entering duty. The number entering duty from the DEP includes some who entered the DEP in the previous year. It does not include those who are still in the DEP at year's end nor those who dropped out without entering duty. See Glossary for abbreviations and acronyms.

A prospect's eligibility to enlist is based on a combination of AFQT score, Aptitude Area Composite scores, and education status. The requirements are:

- High school diploma graduate
  -- AFQT percentile, 16 or higher
  -- Score of 85 in at least one Aptitude Area

- High school equivalency holder
  -- AFQT percentile, 31 or higher
  -- Score of 85 in at least one Aptitude Area

- Non-high school graduate
  -- AFQT percentile, 31 or higher
  -- Scores of 85 in at least two Aptitude Areas

The Later Stages. Applicants who have passed the AFQT hurdle next are examined for their physical soundness. About 22 percent of the male and 44 percent of the female candidates fail this screen, mostly for being overweight.

Each prospect who has made it this far now meets with a uniformed Guidance Counselor, who tries to place the candidate in the most appropriate Army job. Identifying the job openings available at that precise moment is the end product of a complex Army system for forecasting upcoming training needs in each field. The computer-based system (REQUEST) that coordinates the information for assigning training slots is based on minimum qualifications; each MOS has a cutoff based on ASVAB aptitude scores. While a candidate may meet the minimums for many specialties, options are limited by the availability of training slots, the Army's current priority on filling each MOS, and the candidate's own preferences.

If the applicant decides to enlist, he or she accepts an MOS offered by the Guidance Counselor. Generally a contract is signed for a specific MOS, but in some cases the contract is for a general career management field, with the specific MOS being assigned early in training.

The candidate who signs a contract is given a date to report at a Reception Battalion and enters the Delayed Entry Program (DEP) to wait for a training slot, or perhaps to finish high school. The wait averages 90 days, but can be much longer for MOS that are considered very desirable. About 5 percent of these enlistees drop out during the DEP wait. For the remainder, the final step is to report to a Reception Battalion for initial processing.

Controls and Incentives

Constraints on Enlistment. Given a youth manpower pool that numbers in the millions, what constraints bear on the number who will become Army enlistees? It is fair to say that the first is simply competition -- from the other military Services, from higher education and technically oriented training institutions, and from prospects of employment in business and industry. Also, some candidates may be discouraged by the fact that the Army has a considerable number of combat-oriented jobs that do not offer a shadow training in a civilian specialty.

Various controls or influences on who is permitted to enlist have been mentioned. Requirements include:

- Age -- between 17 and 35.

- AFQT Category -- Category level combined with Aptitude Area scores. Applicants in Category V are, by law, not eligible.

- Education -- not by itself, but a higher AFQT score is required of applicants

who lack a regular high school diploma.

- Moral records.

- Physical standards -- tests, with weight and height limitations.

- Regulations -- for example, a cap on the number of dependents one may have.

Constraints on Classification. Why can't the Army give every applicant the job he or she wants? Because the Army may not have a vacancy in training for the applicant's preferred MOS, and the applicant may not meet the aptitude/physical minimums for a job. Also, because the Army takes care to spread high aptitude applicants across all MOS, any given MOS may be temporarily closed to a high-aptitude applicant.

Incentives for Enlistment. One of the hard facts about trying to build a superior fighting force is that selection works both ways. To compete with the other Service and with the private sector for the top-quality young people in the target pool, the Army offers a number of special inducements to get them to select the Army. "Critical skill" bonuses and educational incentives are very effective.

A popular inducement has been the "training of choice" in a specific school training program (if the applicant meets the requirements, and if training slots will be available). In combination with this option, or separately, there may be guaranteed initial assignment to particular commands, units, or locations, mainly in the combat arms or in units that require special technical skills.

In a time of downsizing and tightened recruiting budgets, thoughtful use of incentives helps maintain a flow of highly qualified young people into Army life.

Yes!
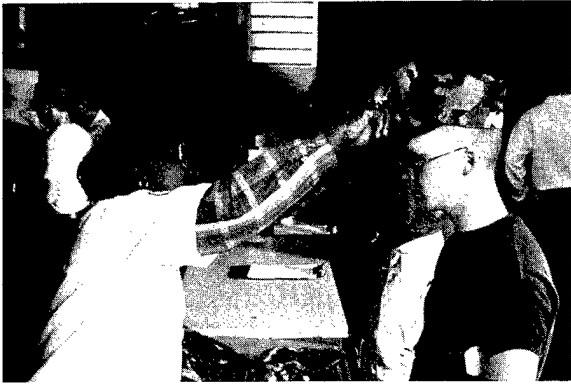
You're
In
The
Army
Now--

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

Military Questions and Research Answers

2

The Army's emphasis on selection and classification research since the early 1980s came about for the varied reasons noted in Chapter 1. The combination of a declining manpower pool and increasing technical demands underscored the need to improve and refine the ways we test and evaluate applicants' potential. The Army needed valid information about human abilities and behaviors as a basis for the continuing drive to improve soldiers' performance.

ARI started its inquiry with these questions:

- *How good is the ASVAB?*

  - Just what does it measure? How well does it foretell success as a soldier? How well do the Aptitude Area Composites work? How should MOS qualifying standards be set?

- *What other types of selection tests might be useful?*

  - How can we match persons with jobs better by measuring additional human traits? Can we cut down on turnover by improving selection? How early can we identify potential leaders?

- *What levels of quality in personnel does the Army need?*

  - What mix of quality levels should be sought? How can we best employ recruits of lower aptitude?

- *How can we learn more from our measurements of job performance?*

- *How can we assure that tests are fair?*

- *How can we judge whether a possible change would be cost-effective?*

What we learned from study of these and related questions forms the subject matter for the remainder of this report.

A Range of Perspectives

The Congressional directive issued in 1981 required the Services to determine how well ASVAB predicts job performance (up to then, research had focused only on ASVAB's power to predict performance in training). DoD's response to this mandate was the Joint-Service Job Performance Measurement (JPM) project, which required each Service to find out how closely individuals' ASVAB scores match up with measures of their subsequent success in military jobs.

As the 1980s began, the Army was laying plans for research into making better use of its personnel. The mandate for the ASVAB project became part of the Army's plans for research on a wider range of issues on the effectiveness of soldiers.

A Long-Term Point of View: Project A/Career Force

Plans evolved to follow a group of soldiers from enlistment, through training, first-tour assignments, and reenlistment decisions, and,

for some, into their second tour.[46] A long-term "longitudinal" research database would be developed, linking new recruits' qualities with their performance in the Army at successive later stages.

The result was a landmark research program, long term and large scale. Project A/Career Force (Figure 2.1) were designed to give the Army the greatest possible increase in overall performance and readiness that could be obtained from improving selection and allocation of enlisted personnel. These sequential projects provided an integrated examination of performance measurement, selection and classification methods, and allocation procedures to meet the multiple goals of managing the Army's human supply.[7]

The combined design made use of two major cohorts of soldiers (new accessions for 1983/84 and for 1986-87), both of which were followed from enlistment into their second tour of duty. The size of the samples was noteworthy; the initial sample for the longitudinal testing of the 1986-87 cohort was well over 45,000 new recruits. The research flow and the size of the various samples as they moved through the Army life cycle are shown in Figure 2.2.

For each sample the entrance "predictor" measures consisted of the pre-enlistment ASVAB and either the Project A Trial Battery of new tests or its later version, the Experimental Battery. The array of measures of success in the Army for each cohort reflected the need to assess the soldiers on many aspects of their performance (a) after training, (b) during first tour, and (c) during second tour. Thus long-term comparisons could be made between what their entrance tests indicated and their actual performance.

Figure 2.1. Project A/Career Force

Project A -- Improving the Selection, Classification, and Utilization of Army Enlisted Personnel
                    1982-1989

Career Force -- Building and Retaining the Career Force: New Procedures for Accessing and Assigning Army Enlisted Personnel
                    1989-1995

Objectives for Project A
- Develop new measures of job performance to use as criteria for validating selection/classification measures.
- Validate existing selection measures (primarily the ASVAB) against both existing and project-developed criteria.
- Develop and validate new selection and classification measures (Experimental Predictor Battery).
- Develop psychology's first-ever model of what performance consists of--this for entry-level Army jobs.

Objectives for Career Force
- Develop measures of second-tour performance as an NCO, based on the Project A prototypes.
- Carry out a predictive incremental validation of (a) the ASVAB and the Project A Experimental Predictor Battery, (b) measures of training success, and (c) the Project A first-tour performance criteria, against the second-tour job performance measures.
- Develop a model of second-tour NCO performance.
- Estimate the degree of differential prediction across (a) major domains of predictor information, such as abilities, temperament, and interests; (b) major factors in job performance; and (c) different types of jobs.

Note. See the Glossary for explanation of some of the technical concepts.

Figure 2.2. Project A/Career Force Research Flow and Samples

| 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 |

8/1/82
File
Data
Analysis

CV = Concurrent Validation
LV = Longitudinal Validation
P = Predictors
T = Training Performance
I = First Tour Performance
II = Second Tour Performance

83/84
Concurrent
Sample
Trial
Predictor
Battery
and
First-Tour
Performance
Measures
N = 9,500

CVI

83/84
Concurrent
Sample
Second-Tour
Performance
Measures
N = 1,000

CVII

86/87 Longitudinal Sample	
Experimental Predictor Battery N = 45,000	End of Training Performance Measures N = 45,000
LV	LVT

86/87
Longitudinal
Sample
First-Tour
Performance
Measures
N = 10,000

LVI

86/87
Longitudinal
Sample
Second-Tour
Performance
Measures
N = 1,500

LVII

| 83 | 84 | 85 | 86 | 87 | 88 | 89 | 90 | 91 | 92 | 93 |

The soldiers assessed were drawn from 21 MOS that had been carefully selected to represent job variations and Army priorities. Nine MOS formed a subgroup that received the full array of tests; for cost-control reasons, the remaining MOS did not receive hands-on and other MOS-specific job performance tests (which are very expensive to develop and administer).

The findings from the many facets of this research program and its spinoffs are presented in Chapters 3-6 of this report.

Doing Research on People in a Military Setting

Research done in a military setting tends to be more applied (i.e., problem solving) than basic (i.e., theoretical). Key aspects in such work are precisely defining the problem and bridging the gap between the different needs of research and military operations.

Defining the Problem

In either theory or practice, civilian or military, solving a problem has to start with correctly defining the problem. What the problem is depends on who is looking at it, and concluding whether research is successful depends on what the consumer needs to do

Figure 2.3. Viewpoints

Take the question: *How can we improve Army selection and classification?*

How would that be answered by:

The recruiter -- who has to screen all those young people?

The test operator -- who has to give and score the tests, pass the results on quickly, and protect the integrity of the information at each stage?

The analyst -- who has to bridge between operations and testing to evaluate the impact of different possible MOS qualifying scores?

The counselor -- who has to sell the qualified applicant on one of the available Army training courses?

The trainer -- who has to take whoever comes along and turn them into competent soldiers?

The unit commander -- who has to merge each new individual into a competent team?

The personnel planner -- who has to fill quotas, meet force budgets, and balance quality?

The operations planner -- who has to produce a force that is combat ready?

The accountant -- who has to figure out how much it costs to make a change, and whether it is worth it?

and

The prospect -- who is interested, perhaps, in becoming a soldier with a good future?

with the results (see Figure 2.3). On top of these operational views of "the problem" are the scientist's, which has to do with figuring out how to collect data and do analyses that really address the problem. These concerns have more to do with the meaning of the research than with running the Army.

Two Different Worlds

If it is to be worth performing, research must abide by its scientific ground rules. The researchers must recognize and spell out assumptions; samples and data sets must be both large enough to provide reliability and representative enough to permit broad application; tests must be administered uniformly if their results are to be trusted, and so on.

The military sponsor of research needs to recognize that such concerns safeguard the integrity of the research data, and thus the power of the research to address the Army's problems. That translates into not only arranging for the mechanics of the research but also adjusting them to meet scientific standards.

On the other hand, scientists must plan their work so that it can fit into the daily routine of an Army that has to continue carrying out its mission.

Project A/Career Force presented many challenges to balance these practical and scientific interests. Figure 2.4 gives some of the details going into the resolution of one such issue: picking the sample of Army MOS to test. Choice of the sample of MOS was crucial if the results of the project were to be applicable to the whole system of Army MOS.

Predicting Army job performance has an obvious prerequisite: an account of what Army job performance is. The next chapter starts presenting the findings of this research with a major contribution to organizational psychology: the first scientifically grounded, comprehensive model of job performance, this model being for Army enlisted jobs.

Figure 2.4. Merging Two Sets of Requirements

<u>Objective</u> (Project A): From the overall list of 276 entry-level MOS in 1984, choose a representative sample of MOS for a long-term project involving predictor and performance testing aimed at improving selection and classification.

Army Requirements	**Research Requirements**
• Produce results that will support conclusions about the Army's Career Management Fields	• Choose MOS so that all ASVAB Aptitude Area (i.e., person-job matching) Composites are represented
• Make sure that Combat MOS are thoroughly covered	• Include MOS from enough different Career Management Fields to permit extrapolation
• Produce results that will hold both for the majority and for women and minority groups	• Choose high-density MOS so that the number of available individuals in each will be large enough to provide valid and reliable results, including those for subgroups and second-tour soldiers
• Evaluate ASVAB's power to distinguish levels of future performance not just in the first enlisted tour, but also into the second tour	

<u>Result</u>: A selection of 21 MOS that meet all of the specifications.

Basic

SOLDIER SELECTION: PAST, PRESENT, AND FUTURE

What Comprises Army Enlisted Job Performance?

3

In selecting people to serve in the Army and in matching them with the right jobs, which human qualities are really useful to know about? The Army's long-range Project A/Building the Career Force research program pursued this question in depth.

To start, this program created the first-ever comprehensive scientific account of a job -- here the Army enlisted job. With that foundation, scientists could evaluate how well the Army's system for selecting enlisted applicants succeeds in distinguishing the ones who will go on to do well from those who will not.

At the heart of this system is the Armed Services Vocational Aptitude Battery (ASVAB). In this research program, a wide range of human attributes that are not measured by ASVAB were also evaluated. In these pages you will find a sampling of the results and conclusions about enlisted performance and the human qualities on which it is based.

The Army is committed to recruiting only the best qualified young men and women. . . . The impact of more than 20 years of success in attracting and retaining quality volunteers has permeated the entire spectrum of Army life.
LTG Thomas P. Carney
Deputy Chief of Staff for Personnel

If these conclusions are not surprising, that could give them the ring of truth. Or it could make you ask "So, what's new?" What's new is that the conclusions have moved from the realm of anecdote and arguable opinion to a solid factual basis. Also new is the creation of tests for measuring more of the whole person and using that expanded view to improve the totality of Army job performance.

Selection and Job Performance

Improving selection is clearly a worthy goal, but how can we tell whether this has happened? Selection amounts to predicting how well an applicant will perform on the job. There are many ways to assess success (attrition rate, for example), but one is basic: an acceptable level of job performance.

But that doesn't take us very far. What levels of performance are acceptable and to whom? How can we measure performance? For that matter, what is performance?

Such questions weighed heavily following the late 1970s when large numbers of low-aptitude youth enlisted as a result of the misnorming of ASVAB. Existing performance measures did not help to relate individuals' ASVAB scores to their success in the Army. Grades given at the end of training had been pretty much replaced by a Pass/Fail system that recycled most failures, helping them to qualify eventually. The Skill Qualification Tests (SQT) of job proficiency varied in quality across MOS. There was no clear link between test scores at enlistment and the fragmented records of a soldier's later success.[46] Furthermore, the situation led to difficulties in the NCO leadership corps, making it evident that we need to predict more than entry-level performance.

Most basic is the question of the "nature of performance." It can be defined as the ability to perform a set of tasks, but when someone *is able* to do the tasks, that does not insure that the person *will* do them. Which human attributes lead one person to perform the tasks easily and another to fail? Does the difference lie in the persons, the tasks, or the efforts to evaluate?

Measuring the Measures

Since the main purpose of the selection process is to identify people who will perform well on the job, a prerequisite to evaluating selection is a decision on how success is going to be measured. Experimental tests are evaluated by how closely persons' scores on them relate to measures of those same persons' performance. A new test could be useful only if persons who score high, middling, or low on it strongly tend to score, respectively, high, middling, or low on measures of their success on the job. Thus, the true starting point for the whole process is how we measure performance.

Exploring the widest possible range of performance measures against which ASVAB (and the planned experimental predictor tests) could be assessed was an early goal in Project A. These surveys would also yield ideas as to what "performance" may include, and would be doubly useful in later work on such questions as the effects of different levels and kinds of aptitudes.

At the outset, Project A assumed that performance involves two basic types of activity. One type is specific to a particular job and reflects technical competence or behavior that is not required for other jobs (i.e., it is MOS-specific). The second type exists in the same way for every job (i.e., it is Army-wide) and includes not only particular

skills, such as first aid, but also such elements as teamwork and persistence.

An array of different ways of evaluating performance was developed in Project A after exhaustive analysis of MOS job content and task criticality. Job sample (hands-on) measures and written job knowledge tests were developed for typical samples (as judged by subject matter experts) of essential job tasks. Rating scales, to be filled out by both supervisors and peers, were set up to assess behavior in both MOS-specific and Army-wide activities. Personnel records and databases were examined to check the informativeness of individuals' histories.

These tests--the criterion measures listed in Figure 3.1--were tried out on a sample of 9,500 first-tour soldiers and then, after revisions, were used to measure performance in the long-range phase of the program. The MOS making up the samples are listed in Figure 3.2. The test array produced more than 150 scores per individual, which were clustered into 22 basic scores of six types: hands-on performance tests, job knowledge tests, Army-wide and MOS-specific ratings, a rating of predicted performance in combat, and administrative performance indicators.

To obtain estimates of NCO potential, the hands-on, job knowledge, and ratings measures were designed so that higher technical skill-level and supervisory job requirements could also be assessed. The increasing importance of the supervisor role relative to technical job requirements led to inclusion of two new types of measures in the second-tour array of performance tests: a written "situational judgment test" and supervisory simulation (role playing) exercises.

Figure 3.1. The Measures of First-Tour Performance in Project A

Ten behavior-based rating scales to measure factors in performance that are not specific to a particular job (i.e., Army-wide):
> Technical Knowledge/Skill
> Peer Leadership and Support
> Demonstrating Effort
> Self-Development
> Maintaining Equipment
> Following Regulations
> Self-Control
> Integrity
> Military Bearing
> Physical Fitness

A single-item rating of overall job performance.

A single-item rating of NCO potential.

A 19-item instrument for rating an individual's expected performance in combat.

MOS-specific rating scales (from 6 to 11 per MOS) to measure factors in job-specific technical and task proficiency.

Job-sample (hands-on) tests of MOS-specific task proficiency (14-17 critical job tasks per MOS).

Paper-and-pencil job knowledge tests to measure both task-specific and common job knowledge (150-200 multiple-choice items representing 30 critical job tasks in each MOS).

Performance indicators from administrative records (via self-report or computerized records):
> Number of Awards and Certificates
> Physical Fitness Qualification
> Number of Disciplinary Infractions
> Promotion Rate

Source: Project A[10]

Beyond establishing the measures of job success to compare with scores on ASVAB and new predictor measures, these diverse ways of looking at performance had other potential uses. They would, for example, provide avenues for fine-tuning MOS classification and job assignment decisions, detecting possible gender or ethnic bias, and comparing cost-effectiveness options.

Procedures for using the various kinds of tests so that reliable performance results are obtained (e.g., training for test administrators to ensure uniform testing conditions; techniques for dealing with missing data) are another product of Project A that is now available for use of the tests in other settings.[10,11] The extensive database obtained in the successive phases of performance testing has also been utilized in other research projects.

What Is Performance?

Overall, the goals for the performance measurement side of Project A were to develop the first-ever general, scientifically grounded description of job performance, this one for entry-level Army jobs, and then to develop reliable and valid measures of each factor in that performance.

This approach was guided by a model that views performance as being "truly multidimensional. There is not one outcome, one factor, or one anything that can be pointed to and labeled as job performance."[10] It is, in fact, a wide variety of behaviors -- things that people do -- that are judged to be important for accomplishing the goals of the organization. The task is to create order from such an assortment.

A series of complex analyses to determine the major factors in the array of first-tour performance measures[6,7,10,34] led finally to five factors, which are defined in Figure 3.3. The factors, termed the dimensions of performance, are:

Figure 3.2. The MOS in the Project A/Career Force Long-Range Program

Full Treatment (Batch A)			
MOS	**Title**	**MOS**	**Title**
11B	Infantryman	71L	Administrative Specialist
13B	Cannon Crewmember	88M	Motor Transport Operator
19E/K	Armor Crewman	91A	Medical Specialist
31C	Radio Operator	95B	Military Police
63B	Light-Wheel Vehicle Mechanic		

Predictor · MOS Specific-Hands-on Job Knowledge · Army-Wide Performance · School Knowledge

Partial Treatment (Batch Z)			
MOS	**Title**	**MOS**	**Title**
12B	Combat Engineer	54B	Chemical Operations Specialist
16S	MANPADS Crewman	55B	Ammunition Specialist
27E	TOW/Dragon Repairer	67N	Utility Helicopter Repairer
29E	Comm.-Electronics Radio Repairer	76Y	Unit Supply Specialist
		94B	Food Service Specialist
51B	Carpentry/Mason Specialist	96B	Intelligence Analyst

Predictor · Army-Wide Performance · School Knowledge

- Core Technical Proficiency
- General Soldiering Proficiency
- Effort and Leadership
- Personal Discipline
- Physical Fitness and Military Bearing

The first two factors, based on hands-on and job knowledge tests of proficiency, reflect, respectively, the technical nature of each MOS and the content that is common across almost all jobs. The remaining factors are based mainly on ratings and administrative records.

The third factor, Effort and Leadership, includes the most scales and the broadest content. The Personal Discipline factor focuses clearly on possible behaviors that have the effect of avoiding trouble. The final factor is quite specific, focusing on military appearance and physical conditioning.

In general, this concept of what enters into the behavior of first-term soldiers fit the data from all of the MOS tested. That is, it was confirmed as an appropriate and reasonable account of the nature of performance.

Over time, the first two factors came to be known as the "Can-Do" dimensions, showing that a soldier is capable of performing these MOS and common duties. The other three factors came to be known as the "Will-Do" dimensions, reflecting a soldier's willingness to perform effectively over time.

Figure 3.3.   The Dimensions of First-Tour Job Performance: Factors and the Basic Scores Defining Them

"Can-Do"

Core Technical Proficiency (CTP) -- How well the individual performs the tasks that are central to the specific MOS. They are the primary definers from job to job.

- Hands-On Test - MOS-Specific Tasks
- Job Knowledge Test -- MOS-Specific Tasks

General Soldiering Proficiency (GSP) -- How well the individual performs common soldiering tasks.

- Hands-On Test - Common Tasks
- Job Knowledge Test -- Common Tasks

Effort and Leadership (ELS) -- How well the individual exerts effort over the full range of tasks, perseveres under adversity, and leads/supports peers.

- Administrative Index - Number of Awards and Certificates
- Army-Wide Overall Effectiveness Rating
- Army-Wide Effort/Leadership Rating
- Average of MOS Ratings

"Will-Do"

Personal Discipline (PD) -- The degree to which the individual adheres to Army regulations and traditions, shows self-control and responsible behavior, and does not have disciplinary problems.

- Administrative Index - Number of Articles 15 and Flag Actions
- Administrative Index - Promotion Rate
- Army-Wide Personal Discipline Rating

Physical Fitness and Military Bearing (PFB) -- How well the individual maintains appropriate military appearance and stays in good physical condition.

- Administrative Index - Physical Readiness Score
- Army-Wide Fitness/Bearing Rating

Attention next turned to the junior NCO positions, which are so critical for the continuing effectiveness of the Army. Study of documentary and operational second-tour sources verified that the technical parts of performance as an NCO retain the same type of content, but with more complexity and expertise. Leadership and supervisory activities were added to the performance measures. A sixth dimension--Leadership--emerged for the second-tour performance model; the first-tour dimension termed Effort and Leadership was modified as Effort and Achievement for the second tour.

The entry-level and junior NCO performance models derived from this in-depth Army research have formed the basis for a more generic performance model, applicable to all military and civilian jobs in the labor force. This model promises to be influential in the larger field of personnel research.[8]

How Good Are ASVAB and the New Project A Tests?

4

Little could be done to assess ASVAB's actual accuracy in predicting performance until the factors in job performance were isolated and ways of measuring them were devised (Ch. 3). Now we are ready to see what was learned about the power of ASVAB and the new tests to play their intended role in maintaining a high level of quality in Army personnel and performance.

How Good is the ASVAB?

That simple question was the starting point for the expanding sequence of selection and classification research during the past 15 years. The Congressional mandate to provide proof that ASVAB predicts later job performance reinforced trends already underway to examine selection closely.

The power of the ASVAB--in use since 1976 as the main selection instrument for the Armed Forces--to predict enlistees' later performance had previously been tested (validated) only against their performance in training. Far more difficult is the long process of following soldiers over the years to see how well the ASVAB predicts their success in first- and then second-tour jobs.

Prediction of Performance

Once these dimensions were identified, it became clear that the relationship between predictor -- the ASVAB in this case -- and performance depended very much on which dimension of performance was being examined.[2] (See Figure 4.1.)

The ASVAB proved to do very well in predicting both job-specific and general soldiering proficiency of first-term soldiers on the job as well as in school. These outcomes are the "Can-Do" dimensions of performance. Correlating soldiers' predictor scores (ASVAB) with measures of their first-tour performance, the values were .63 for Core Technical Proficiency and .65 for General Soldiering Proficiency. In this kind of research, values as low as .30 are usually considered good. The relation of General Soldiering Proficiency to AFQT Categories is shown in Table 4.1.

**Figure 4.1. Prediction of Job Performance Using ASVAB**
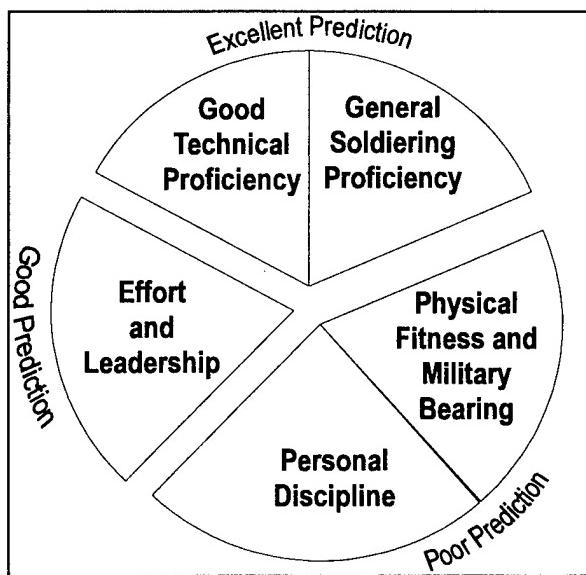
Table 4.1. Prediction of General Soldiering Proficiency in the First Tour Using ASVAB

AFQT Category	GSP Performance Level		
	Superior	Avg.	Below Avg.
I - IIIA	42%	33%	25%
IIIB	26%	36%	38%
IVA - IVB	17%	30%	53%

The ASVAB did less well in predicting the three "Will-Do" dimensions. Here the correlations with performance were lower: .38 for Effort and Leadership (which is still very good), .22 for Personal Discipline, and .21 for Physical Fitness/Military Bearing.[10]

What Does the ASVAB Measure?

Looking at the set of ten cognitive (e.g., verbal, math) and vocationally oriented ASVAB subtests, analysts found four factors: Verbal Aptitude, Quantitative (Mathematical) Aptitude, Technical Knowledge, and Perceptual Speed. The Quantitative factor, for example, is made up of the Math Knowledge and Arithmetic Reasoning subtests. These factors define the specific aptitudes sampled by the ASVAB. Further analysis showed that the Quantitative and Technical factors contribute the most to predicting General Soldiering Proficiency, and the Verbal score is more prominent in predicting Core Technical Proficiency.[10]

The Army wanted to know not only how predictive these tests were but also what they did not measure -- and thus what kinds of new tests might be used to supplement the ASVAB. Review of the literature, plus a lot of experimenting, led to three types of new measures that appeared most promising.[39]

- Temperament/interest/biodata -- information on personality, vocational interests, and life experiences.

- Spatial -- the ability to use information on direction, form, orientation, position, etc., of objects in space.

- Perceptual/psychomotor -- measures of such characteristics as eye-hand coordination.

Development of experimental predictor tests in these additional areas is described later in the section on "Expanding Ways of Predicting Success (and Failure)."

> The question of whether the Armed Services Vocational Aptitude battery (ASVAB) does or does not predict job performance (in addition to training performance) has been answered definitively, in the affirmative. The Army and the Department of Defense are now in firmer position to support their quality goals. In addition, it is now known what aspects of performance could be predicted better with other types of selection instruments.
>
> *-- Final Report on Project A*[10]

ASVAB for Classification

The initial link of a potential recruit with a specific Army job is provided by the ASVAB subtests during the enlistment process. The subtests are grouped into ten Aptitude Areas for MOS that have similar requirements. For example, the General Maintenance Aptitude Area Composite includes General Science, Electronics Information, Auto and Shop Information, and Math Knowledge.

Aptitude Area composites are an ongoing target of research because recent gains in computational power now make it possible to

use the ASVAB subtests and new tests from Project A to optimize performance more effectively in matching persons with jobs. In the early stages of Project A, review of these composites resulted in a number of shifts that improved prediction.

Each MOS has its own cutoff scores on its Aptitude Area composite. An applicant who scores above the line is eligible for that MOS; one who scores below it will have to go elsewhere. In effect, these levels set the standards for an MOS. How and where to establish these standards is a major concern in the effort to choose and use soldiers effectively. Supplementing the present testing information about an applicant's potential in order to improve the job/person match is one of the continuing objectives in selection and classification research.

Expanding Ways of Predicting Success (and Failure)

Trying to get a fix on potential Army recruits beyond that supplied by the ASVAB is hardly a new undertaking. Over the years the Army has tried out--with mixed results--a variety of techniques, especially with applicants who are not high school graduates. More than ever under current conditions, the Army needs efficient recruiting. A sharper focus on predicting the potential job performance of the young people who step into a recruiting office will help to cut down on costly attrition losses, and to match the range of what applicants can do with the range of the complex demands of Army jobs.

Project A's Experimental Predictor Tests

A "touch-all-bases" effort to explore many types of predictors in addition to the cognitive tests of the ASVAB was a major part of the Project A design. The idea was to identify a very wide pool of human attributes that may

affect performance in entry-level MOS, select the most promising, develop measures, and then try out these new tests in pilot and field tests. The winners would be used in the predictor battery in the formal Project A long-range validation.

The hunt began with an in-depth search of the entire personnel selection literature--both cognitive, such as verbal/math and psychomotor abilities, and noncognitive, such as personality and life experiences. An initial list of several hundred possibilities was eventually reduced to these groupings: cognitive abilities, visualization/spatial, information processing, mechanical, psychomotor, social skills, vigor (energy, self-esteem), and maturation/stability.

The next step was to reach agreement on what candidates would best supplement the ASVAB in predicting job performance across all MOS. Many kinds of measures -- cognitive and noncognitive paper-and-pencil tests and inventories, computer-administered tests, archival records -- were developed and tried out. Four main types survived: spatial, perceptual/psychomotor, temperament, and interests.

Finally, in Project A's Concurrent Validation in 1985, the surviving tests were evaluated against the array of job performance criterion measures described in Chapter 3. In this phase of the project, first-tour soldiers took the new Project A predictor tests and the performance measures in the same two-day period. The results provided the raw material for developing the model of job performance -- that is, for identifying the five dimensions, described earlier, that depict the performance of first-tour soldiers.

Once these dimensions were identified, the task was to see how well the different types of

predictive measures predicted them. The broad findings are summarized in Table 4.2, comparing prediction of performance from ASVAB alone with results when each type of experimental test is added to the ASVAB. In this table, it can be seen that the addition of spatial tests produced small but consistent gains in predictive accuracy for each dimension. So, to a slightly lesser extent, did adding the perceptual/psychomotor tests. Noteworthy was the sharp increase that adding the temperament measures produced for the dimensions of Effort and Leadership, Personal Discipline, and Physical Fitness. For example, on Effort and Leadership the predictive accuracy for ASVAB alone was .31; addition of the temperament measures raised the level to .42. Similarly, adding vocational interest measures produced good gains in Effort and Leadership, and Physical

Fitness, as well as lesser gains for the other dimensions.

The spatial/psychomotor and the temperament/biodata tests are described in following sections. Detailed results by individual predictor and performance measures and by MOS are available in many reports (see References).

The product of this developmental marathon was the Experimental Predictor Battery, listed in Figure 4.2, which was used in the long-term -- longitudinal validation -- portion of Project A/Career Force in 1986-87. The predictor battery was administered to more than 45,000 enlistees in 21 MOS entering the Army in 1986-87; about 38,000 had complete predictor data and became part of the long-term assessment process.

Table 4.2. Gain in Predictive Power When Experimental Tests Are Added to ASVAB

Performance Dimension	ASVAB Factors Alone	ASVAB Plus:			
		Spatial Tests	Perceptual/ Psychomotor Tests	Temperament Measures	Vocational Interests
Core Technical Proficiency	.63	.65	.64	.63	.64
General Soldiering Proficiency	.65	.68	.67	.66	.66
Effort and Leadership	.31	.32	.32	.42	.35
Personal Discipline	.16	.17	.17	.35	.19
Physical Fitness and Military Bearing	.20	.22	.22	.41	.24

Note. Nine Batch A MOS, Project A Concurrent Validation correlations.
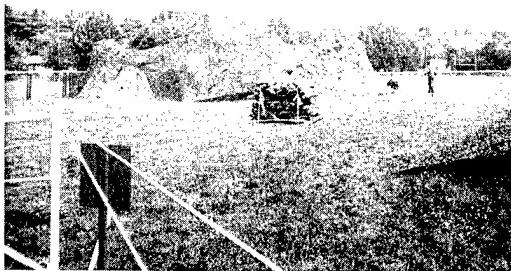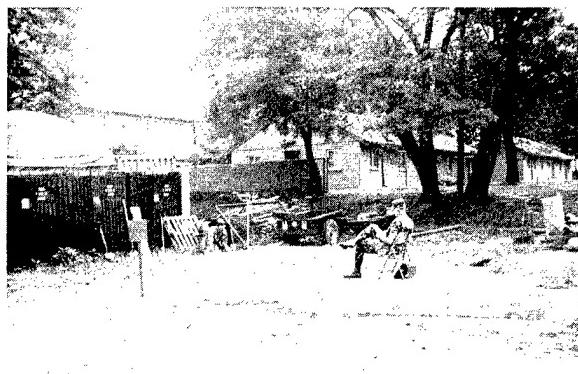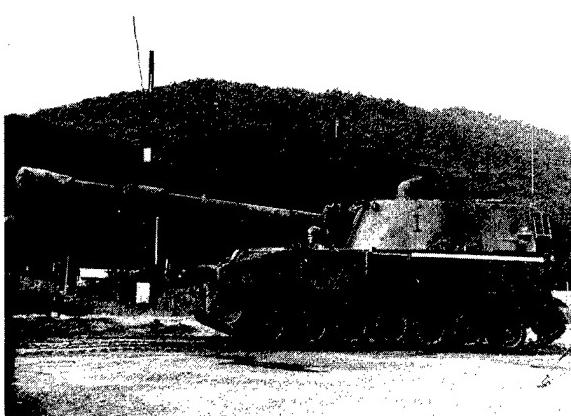
**Testing,
Testing,
Testing,
Testing...**

Figure 4.2. Project A's Experimental Predictor Battery

| Test Type	Name of Measure	Characteristic Tested	
Spatial (Paper-and-Pencil)	Assembling Objects Object Rotation Maze Orientation Map Reasoning	Spatial Visualization-Rotation Spatial Visualization-Rotation Spatial Visualization-Scanning Spatial Orientation Spatial Orientation Induction	
Perceptual/Psychomotor (Computer-Administered)	Simple Reaction Time Choice Reaction Time Short-Term Memory Perceptual Speed/Accuracy Target Identification One-Hand Target Tracking Target Shoot Two-Hand Target Tracking Number Memory Cannon Shoot	Reaction Time (Processing Efficiency) Reaction Time (Processing Efficiency) Short-Term Memory Perceptual Speed/Accuracy Perceptual Speed/Accuracy Psychomotor Precision Psychomotor Precision Multilimb Coordination Number Operations Movement Judgment	
Temperament, Interests, Biodata (Paper-and-Pencil)	Assessment of Background and Life Experiences (ABLE)	Adjustment Dependability Achievement Physical Condition	Leadership Locus of Control Agreeableness
	Army Vocational Interest Career Examination (AVOICE)	Realistic Conventional Social Artistic	Investigative Enterprising

Performance was tested and analyzed at the end of training, within first tour, and within second tour -- as much as six years later. The results are summarized in the sections that follow.

Spatial and Perceptual/Psychomotor Abilities

As noted earlier, the spatial and, to a lesser extent, perceptual/psychomotor, measures tended to show consistent, if not large, gains in predictive accuracy when they were used to supplement the ASVAB in Project A/Career Force.

The spatial tests were of the traditional paper-and-pencil type. They measured individuals' ability to use information about space--for example: orientation, compass direction, location, form, and paths. The perceptual/psychomotor measures, which were computer-administered, tested such abilities as reaction time, perceptual speed and accuracy, and one and two-hand coordination and precision in tracking.

When individual tests were considered, results suggested that these predictors might be especially good at identifying aptitudes for visually guided gunnery. The Army then conducted further research along these lines. In the Skills Selection and Sustainment Program (S3) of the U.S. Training and Doctrine Command in 1988-89, anti-tank weapons crew and tank crew trainees were tested at Fort Benning and Fort Knox, respectively.[51]

Ability in two-hand tracking was found to be the most effective supplement to ASVAB, doubling accuracy in predicting anti-tank gunnery and improving prediction of tank gunnery by a half. To equal the hit rate of students with psychomotor skills in the top one-third, 11 percent more tanks manned by gunners with average psychomotor abilities would be needed.

Possible predictive increases by adding Project A spatial and psychomotor tests to the ASVAB were investigated by ARI in further analyses of Project A data. The new analyses emphasized the effects of individual tests and checked the predictive data against additional criteria, such as SQT scores.[4,5] Including scores from the Project A tests substantially improved the prediction of many criteria, including General Soldiering Proficiency; for some criteria, such as the SQT, improvement was more specific to MOS. In general, spatial and psychomotor tests were approximately equal in improving prediction. The tests that were most useful were Assembling Objects, Reasoning, Map, and Target Identification.[5]

Concurrently with the Army's Project A/ Career Force program, DoD was conducting research into computerizing the ASVAB, which is used by all the Services. Changing to computer delivery permits a more radical advance in group testing: tailored, or *adaptive*, testing. On an adaptive test, the computer scores each answer instantly. The software then selects each next item from a large, well-scaled item pool to be close to the examinee's estimated level of aptitude up to that point.

This test strategy allows reliable and valid measurement with fewer items and in a shorter time than does traditional paper-and-pencil testing. With computers and the newly saved time available for testing, it has become

feasible to supplement ASVAB with spatial and psychomotor tests.

In 1989 DoD set out to conduct a Joint Service tryout of the Services' most promising new computerized tests. This Enhanced Computer-Administered Test (ECAT) project, based in part on the research reported above, selected six of the new tests from Project A for tryout in its battery of nine tests. Because of the high promise of the Project A spatial tests, three of them were converted from the pencil-and-paper medium for ECAT. The practical problem for ECAT was to determine which subset of the 19 ASVAB and ECAT subtests has the most promise to be the operational test battery of the future.

For identifying optimal *batteries* of ECAT subtests, the analyses were carried out on subsets of the 19 tests in the trials. All possible combinations of the tests were tried (yes, thousands of potential batteries) with these two limitations. First, the power of general aptitude to predict "Can-Do" performance is known to be so strong that all of the batteries were designed to include a measure of general aptitude as a core: the Arithmetic Reasoning and Word Knowledge subtests of the ASVAB. Second, batteries were subject to some limits on testing time.

ARI analyzed scores on all 19 predictors for the 11,700 Army, Navy, and Air Force trainees in the sample, along with measures of their end-of-training (mostly hands-on) performance. The batteries were evaluated for effectiveness on four criteria: improving selection, improving person-job matching, and minimizing differences in scores of gender and racial or ethnic subgroups (to maximize fairness).[32]

Sample Tests

Spatial

Object Rotation -- The test has 90 items and a 7½-minute time limit. The subject examines a given object and decides whether an object shown below is a rotation of the given one or, instead, a rotated mirror image. Every given object has five test objects below it, each one requiring a response of "same" or "not same."

Psychomotor

Target identification -- The subject sees a target object above three test objects -- pictures of military vehicles or aircraft. The target object is the same as one of the test objects, except that it may be rotated. In 36 trials the subject must decide which test object is the same as the target object and then press a button denoting that choice. The average response time is calculated across all trials where the subject makes a correct response.

Analysis showed considerable overlap between the batteries along with clear indications that ECAT measures three abilities beyond those measured by ASVAB: Working Memory, Spatial Ability, and Psychomotor Ability.[1] As to the "best battery," no one battery emerged. The group of subtests that ranked highest for one criterion (e.g., effectiveness in selection) did not necessarily rank the same way for the others (e.g., minimizing male-female differences in scores).

A similar comparison of possible test battery combinations was part of Project A/ Career Force, with a sample of 17,000 soldiers from the nine Batch A MOS. The measures used were Technical/School performance (an end-of-training measure), and two on-the-job criteria, Core Technical Proficiency and hands-on test performance. As in ECAT, the Arithmetic Reasoning and Word Knowledge ASVAB subtests were included in all batteries; the other tests used in varying combinations were the other eight ASVAB subtests, the six Project A spatial and psychomotor subtests that are on the ECAT list, and a Project A subtest of Short-Term Memory.[33] An example of the kind of information that can be assembled is given in

Table 4.3, showing the composites most often found in equations for nine MOS to predict first-tour technical proficiency.[37]

As in the ECAT analyses, here choosing the best possible battery requires setting priorities -- for example, is selection efficiency (based on general aptitude) more important than classification efficiency (based on more specific aptitudes)? -- and then making tradeoffs.

These analyses yielded a great deal of information about the effects of balancing different personal characteristics against

Table 4.3. Composites Most Often Found to Predict First-Tour Technical Proficiency

Composite	Source	Times Entered in 9 MOS Prediction Equations
Spatial	Spatial	9
Technical	ASVAB	8
Quantitative	ASVAB	7
Verbal	ASVAB	3
Dependability	ABLE	2
Short-Term Memory	Perceptual	2
Speed	ASVAB	2

different ways of judging performance and different ability requirements in specific jobs.

Dispositions as Predictors: Temperament and Biodata

The "people will be people" truism has led the Services, more than once, to try to make use of biographical information and life experiences (biodata), as well as traits of personality, as indicators of future military success. Biodata and personality measures are used with confidence in the civilian sector, but these tools have a peculiarly strong military appeal; the Services are trying to select promising volunteers not just for a job or a career field but for a unique style of life.

Who, among those thousands of young people, will work well in the structure, the discipline, the unpredictability and risk of the soldier's life? And who, in that aura of discipline, danger, and teamwork, would show something more -- the whatever-it-takes to grow into the kind of soldier on whom everything in the Army is built, the noncom? More than a measure of ability is needed to find such people in that crowd of untried youngsters.

When Project A was undertaken in 1982, temperament measures were generally regarded as poor predictors of performance on the job. Project A staff, however, felt that assessment had often taken too global an approach, not distinguishing between different aspects of performance, and that strong relationships would be found when the focus was on more specific attributes.

Building on extensive literature reviews, testing, and analysis, they put together the Assessment of Background and Life Experiences (ABLE), sampling the individual's behaviors, beliefs, and attitudes

that were thought to indicate underlying traits. In its final form the measure consisted of ten temperament scales and a biodata scale for physical condition (see Table 4.4 for the list of scales). There were also four scales designed to control the quality of the data by detecting careless or dishonest patterns of responding. The total inventory contained almost 200 items.

So what did these dozens of individual items in the various scales mean in terms of how real people live and work? Project scientists grouped the specific scales into the more general qualities that make up what we think of as temperament or character. Thus the attribute of Dependability emerged from the scales indicating Traditional Values, Nondelinquency, and Conscientiousness, and Achievement has its origin in the scales for Work Orientation, Self-Esteem, and Energy Level. The resulting attributes (termed constructs) are listed in Table 4.4, along with a few words of description.

As mentioned earlier, ABLE was part of the Project A Experimental Predictor Battery. Eventual comparison of its predictions with

later performance showed that ABLE added little to ASVAB in predicting "Can-Do" factors, but added more in predicting "Will-Do" factors (Figure 4.3). Dependability scores had the highest relationship with Personal Discipline performance; soldiers in the lowest third on ABLE Nondelinquency scale had 66 percent more Articles 15 and Flag Actions than those in the top third. By contrast, ASVAB does not differentiate among individuals regarding discipline.[53]

Overall, temperament composites had their highest relationship with measures of leadership, job effort, and personal discipline and were not as good at predicting technical proficiency. The ABLE scores have a near zero relationship with ASVAB and educational level, the main measures the Services use in pre-enlistment screening. Therefore, almost all of ABLE's ability to predict job performance is in addition to ASVAB and level of education as a predictor of performance on the job.[30]

A short form of ABLE, developed for the Adaptability Screening Profile project, can provide information that supplements, rather than supplants, typical selection methods. This 70-item version could be useful not only in the military but in public and private sectors.[53] It concentrates mainly on the Achievement, Dependability, and Adjustment attributes.

Researchers have examined data from ABLE to explore other possible applications -- in one case, to see whether it might indicate a tendency to have accidents. Records of at-fault ground accidents for the soldiers in the Project A longitudinal sample were obtained from the U.S. Army Safety Center. Overall, 4 percent of these soldiers were responsible for causing an accident during their first three years of service. It turned out that these soldiers' ABLE scores on Adjustment,

Figure 4.3. Different Predictive Areas for ASVAB and ABLE
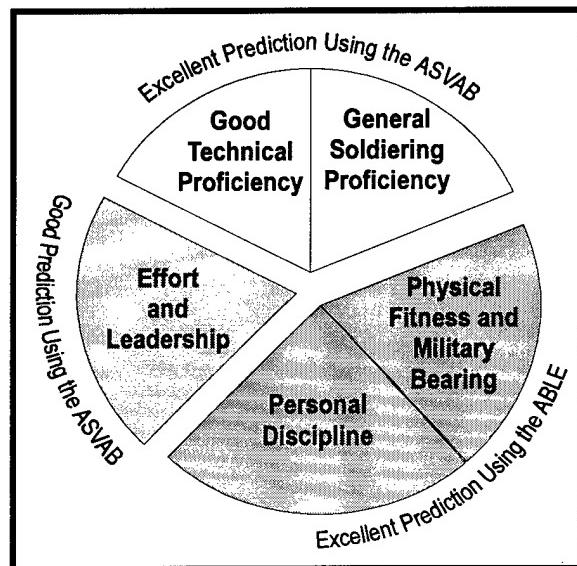
Dependability, and Achievement were predictive of a tendency to be involved in accidents.[26] For example, Infantrymen classed as least dependable (lowest 20%) caused 69 percent more accidents than those classed in the highest 20 percent. In contrast, the ASVAB did not differentiate among personnel causing accidents.

The ABLE was also used for several years to select Marine Security Guards (MSG), who provide security services at U.S. diplomatic and consular facilities throughout the world. Concerns over these Guards' qualifications had mounted after one of them admitted providing information to the Soviet Union while serving in Moscow. Most of the ABLE attributes proved to be related to success in MSG training and performance in the field, which led to this application.[57]

In other instances, ABLE scales, with their detailed and specialized information, were predictive of successful completion of the Ranger course[52] and the course for selecting volunteers to enter Special Forces Training.[12]

Table 4.4. ABLE Temperament Scales and What They Tell Us

Scale	Temperament Constructs (Rational Composite)
Dominance	Leadership Potential: The tendency to enjoy positions of leadership and influence over others.
Work Orientation Self-Esteem Energy Level	Achievement: The tendency to strive energetically for competence in one's work.
Traditional Values Nondelinquency Conscientiousness	Dependability: The tendency to be disciplined, obey and respect rules and regulations, and accept authority.
Emotional Stability	Adjustment: The tendency to be even and positive and to perform well under stress.
Cooperativeness	Agreeableness: The tendency to be pleasant in interpersonal relationships--easy to get along with and a team player.
Internal Control	Locus of Control: The tendency to perceive reinforcements as being under one's control.
Physical Condition	Physical Condition

In the latter case, ABLE scores were related to the reasons for attrition (i.e., voluntary or involuntary). The candidates with very low ABLE scores (74 out of 1,023) had a disproportionately high rate of voluntary withdrawal from the program compared to candidates with higher ABLE scores.[31]

Another questionnaire-type test that was developed in Project A was the Army Vocational Interests Career Examination (AVOICE), an adaptation of an interest inventory from the Air Force. The version included in the Experimental Predictor Battery consisted of 182 items grouped into 22 scales. Respondents were asked to rate how much they would like a variety of jobs, work tasks, spare time activities, and desired learning experiences. The AVOICE scales were grouped into eight composites that indicated these interests:

Rugged/Outdoors	Protective Services
Structural/Machines	Audiovisual Arts
Administrative	Food Service
Social	Skilled Technical

The initial AVOICE analyses indicated that vocational interests predict technical performance and, to the extent that technical performance predicts leadership, interests predict leadership as well. This was particularly evident for the Rugged/Outdoors composite.

A Comparison

In Figure 4.4 we see a comparison of the predictive strength of these different categories of human attributes as found in the Longitudinal Validation. Note here that the ASVAB scores were from the soldiers' days as applicants, before enlistment, and that the new Project A predictor tests were taken during reception processing, just before examinees entered Basic Training. The aspect of first-tour job success reflected in this figure is the

Core Technical Proficiency dimension in the model of first-tour performance (Figure 3.3).[37] This figure shows the pre-eminent power of ASVAB to predict aspects of skilled performance. In the next chapter, the contribution of ASVAB and the other human attributes to the "Can-Do," "Will-Do," and Leadership aspects of performance in the first and second tours will be explored in more detail.

Figure 4.4. Relative Strength in Predicting First-Tour Technical Proficiency

Correlation With Core Technical Proficiency
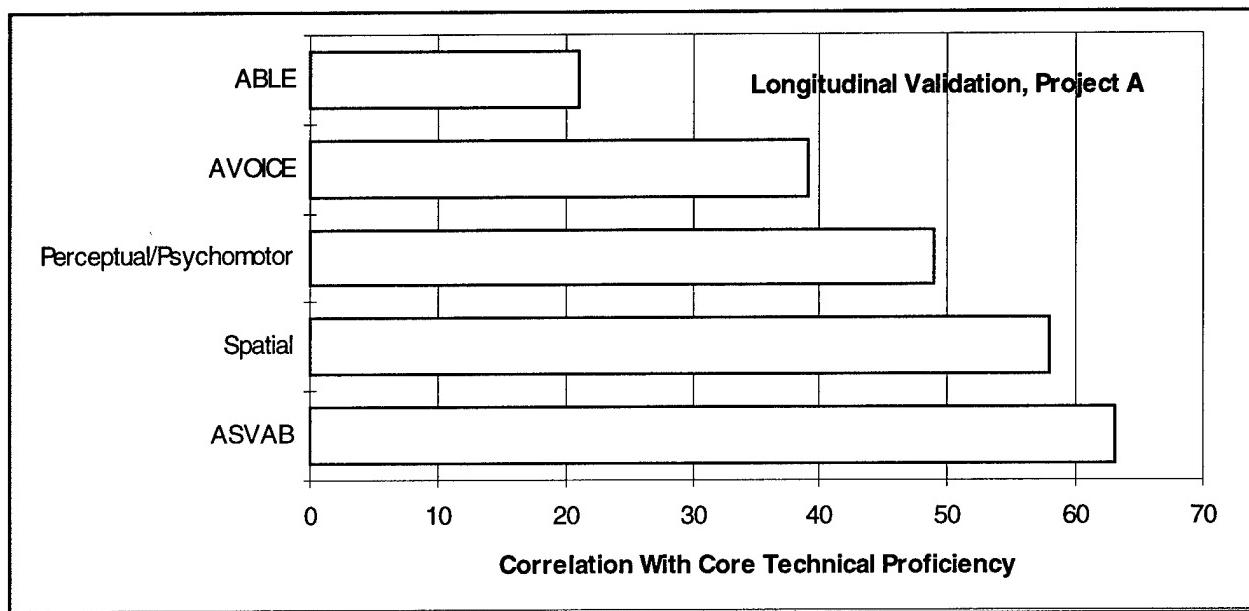
Predicting "Can-Do," "Will-Do," and Second-Tour Performance

Up to this point, we have focused on research relating to enlistment and entry-level standards and performance. In this chapter we summarize some of the findings on predicting the different dimensions of performance, including leadership, in the first and second tours.

Some Findings About Prediction and Performance

Two basic questions about performance -- how to measure and how to predict it -- have come together in Project A and related projects. Thanks to this exhaustive, long-term program, we now have a clear idea of the major factors making up entry-level Army performance and how they can be measured and predicted -- and we have done a lot of measuring.

What Have We Learned About Proficiency?

Starting from the belief that performance is not one thing but many, we have recognized two main aspects -- the soldier's <u>skill</u> in performing certain specified duties, and the soldier's <u>willingness</u> to perform those duties over time. As described earlier, these concepts can be translated into the "Can-Do" factors -- General Soldiering Proficiency and Core Technical Proficiency -- and the "Will-Do" factors -- Effort and Leadership, Personal Discipline, and Physical Fitness. Given this revealing picture of the components of performance, various types of predictor tests administered early in the selection-classification process can be used to forecast an individual's behavior in the equally varied aspects of soldiering performance.

The two "Can-Do" dimensions reflect the dual nature of a soldier's duties: specific technical skills for a specific MOS, and general soldierly skills that are needed in all MOS. In this research, proficiency was measured on the hard facts of daily Army life -- by hands-on and written tests of specific MOS skills on the one hand, and hands-on and written tests of Army-wide skills on the other. A comparison of results from these first tour-of-duty tests with the expectations from the wide-ranging predictor tests some two years earlier yielded many insights. For example:

- The ASVAB factor scores are the best predictors of proficiency on the job. And they are very good at it: The statistical relationship (i.e., correlation) between forecast and first-term performance is .65 for General Soldiering Proficiency and .63 for Core Technical Proficiency.

- The spatial ability and the perceptual-psychomotor composites provide substantial prediction of both the proficiency dimensions.

- However, adding the spatial and perceptual-psychomotor tests to the ASVAB tests produces only a small gain in predicting proficiency. The

power of ASVAB alone was found to be so high that there is not much room for improvement there.[30] In any event, for both proficiency dimensions the spatial domain is the new predictor having the most effect, over and above that provided by ASVAB.

- Among ASVAB scores the Quantitative and Technical composites contribute the most to predicting General Soldiering Proficiency. The Verbal composite plays a more prominent role in predicting Core Technical Proficiency.

- Vocational interests predict technical performance well.

- In later comparisons to identify a "best" predictor battery, the list of tests that had the greatest success varied considerably by MOS and by the way in which success was defined (e.g.,

more efficient classification, lower subgroup differences). Table 5.1 provides an example of the ten best predictors of Core Technical Proficiency for three MOS.

What Have We Learned About Personal Discipline?

The soldier who gets into trouble is a major headache for everyone around him who is trying to build a proficient military unit. Disruptive to training or operations, costly in terms of punishment or attrition, such a soldier is a handicap and a waste. The most efficient way to handle such persons is to keep them out of the Army in the first place. The ASVAB gathers little information bearing on this problem, and this is a major argument for including biodata and temperament predictors in pre-enlistment screening. It is hardly surprising that Personal Discipline emerged as one of the three "Will-Do" dimensions of military performance.

Table 5.1. Sample Test Lists for Best Selection to Predict Core Technical Proficiency

MOS 11B Infantryman	MOS 88M Motor Transport Operator	MOS 91A Medical Specialist
ASVAB Technical	Spatial	ASVAB Technical
Spatial	Structural/Machine Interests	Spatial
Movement Time	ASVAB Quantitative	ASVAB Quantitative
ASVAB Speed	ASVAB Technical	ASVAB Speed
Number Speed/Accuracy	Short-Term Memory	ASVAB Verbal
ASVAB Quantitative	Food Service Interests*	Short-Term Memory
ASVAB Verbal	Rugged/Outdoors Interests*	Rugged/Outdoors Interests
Rugged/Outdoors Interests	Dependability	Perceptual Accuracy
Achievement	Cooperativeness	Dependability
Adjustment	Psychomotor	Psychomotor*

* Indicates that the relationship was negative.

Personal Discipline is defined as the degree to which the individual "adheres to Army regulations and traditions, shows self-control and responsible behavior, and does not have disciplinary problems." In Project A, the three measures that are the markers for this dimension are a disciplinary index (Number of Articles 15 and Flag Actions), the soldier's rate of promotion, and supervisor and peer ratings on the Army-Wide Personal Discipline scale.

Because Personal Discipline is an aspect of performance that is under motivational control, theoretically it should be predicted best by personality or interest measures, not by aptitude or skill. In fact, the Dependability attribute in the ABLE temperament battery has the highest relationship with the Personal Discipline factor. The scales that make up Dependability are Conscientiousness, Traditional Values, and Nondelinquency; soldiers scoring in the lower third on Nondelinquency had 66 percent more Articles 15 and Flag Actions than those scoring in the upper third on this measure.[61] Vocational interests showed little relationship to Personal Discipline.

What Have We Learned About Physical Fitness?

While physical standards for enlistment were not part of the Project A agenda, the general importance of physical capabilities in accomplishing the Army's mission was evident from the identification of Physical Fitness and Military Bearing as one of the three "Will-Do" dimensions of entry-level performance. This performance factor was measured by a combination of a soldier's Physical Readiness Test Score (from Army archives) and the supervisor and peer ratings of the soldier's physical condition and military bearing.

The most successful predictor of physical performance was the ABLE Fitness composite. In that composite, the heavy hitter was the one biodata component of ABLE, the Physical Condition scale, which inquires about physical activities and conditioning. ASVAB provides little predictive information in this regard, and the interest predictors (AVOICE) also had little effect.

The Dependability Scales --

Conscientiousness. Assesses a person's tendency to be reliable. The person who scores high on this scale is well organized, planful, prefers order, thinks before acting, and holds him- or herself accountable. The person who scores low tends to be careless and disorganized, and acts on the spur of the moment.

Traditional Values. Assesses a person's acceptance of societal values. The person who scores high on this scale accepts and respects authority and the value of discipline. The person who scores low is unconventional or radical and questions authority and other established norms, beliefs, and values.

Nondelinquency. Assesses a person's acceptance of laws and regulations. The person who scores high on this scale is rule abiding, avoids trouble, and is trustworthy and wholesome. The person who scores low is rebellious, contemptuous of laws and regulations, and neglectful of duty or obligation.

What Have We Learned About Enlisted Leadership?

Tracing the thread of growing enlisted leadership capability -- from glimpses of leader potential in that pre-enlistment screen, through maturing steps during first-tour duty, and into the reward stage where a competent sergeant takes on the next generation -- was one of the main aspects of the Project A/Career Force long-range research sequence. The Career Force phase was designed to study the nature of leadership by watching it develop over time. This tracking of leader development was a major part of the research on Army job performance; good noncoms make a high impact on the Army.

For the first tour, the "Will-Do" dimension of performance that begins the maturing process was called Effort and Leadership. For the second tour, when the jobs call for greater technical expertise and supervisory

responsibility, the dimension was redefined as Effort and Achievement and expanded to give rise to a sixth dimension, Leadership. The tasks for the other performance dimensions -- General Soldiering and Core Technical Proficiency, Personal Discipline, and Physical Fitness -- were also made more advanced for second-tour soldiers.

The measures identifying first-tour Effort and Leadership were diverse, consisting of Number of Awards and Certificates (an archival index), supervisor and peer ratings of overall effectiveness and of effort/leadership, and an average of the MOS-specific ratings. The best predictor of this performance factor was, to the surprise of many, ASVAB. Not only does ASVAB predict an individual's own proficiency on the job in the Army, but also that person's capacity for bringing out good performance in others. It was the Technical factor of ASVAB that best predicted this outcome, but the best prediction occurred when all of the various predictor factors were included in the calculation.

The Effort and Leadership dimension actually bridged between the "Can-Do" and "Will-Do" aspects of job performance. It is also strongly predicted by the Work Orientation, Self-Esteem, and Energy Level scales of ABLE. Clearly, the leadership aspect of performance has many facets.

The research focus then moved to diagnosing what is involved in leadership at the second-tour level (see Table 5.2 for the content of the expanded behavioral ratings). The measures developed for the Effort and Achievement dimension were Awards and Certificates, ratings of technical skill and effort, an overall MOS performance rating, and ratings of overall effectiveness and predicted performance in combat. The

Example of Supervisory/Leadership Performance Ratings

ACTING AS A ROLE MODEL FOR SUBORDINATES		
Motivates subordinates to perform effectively through personal example, including demonstrating high standards of military appearance, bearing, and courtesy; is a model supervisor for subordinates to look up to by demonstrating exemplary behavior as a soldier.		
Falls below standards and expectations for performance in the category "Acting as a Model" compared to soldiers at same experience level.	Meets standards and expectations for performance in the category "Acting as Model" compared to soldiers at same experience level.	Exceeds standards and expectations for performance in the category "Acting as a Model" compared to soldiers at same experience level.
1 2	3 4 5	6 7

Leadership dimension measures were promotion rate, Army-wide ratings of the leading/supervising composite, and scores from two new types of tests: the Situational Judgment Test, in which the soldier selects most and least appropriate behaviors in typical supervisory situations, and the Supervisory Simulation Exercises, in which an evaluator plays the part of a subordinate to be counseled or trained by the examinee. The analyses provided a foundation for such inquiries as trying to predict second-tour potential during the pre-enlistment screening, and setting the standards for reenlistment.

Looking Toward a Second Term

What qualities does the Army want in its noncoms? That query translates into an operational question: What information in soldiers' records would be useful in managing reenlistment? And that raises still another question that goes back in time: Can the Army identify, back at initial selection for enlistment, applicants who are likely to succeed not just in the first tour but also as career soldiers? A crystal ball that worked would be a nice piece of equipment for pre-enlistment screening.

The Army must decide which individuals, among those willing to enlist or reenlist, can handle the rigors and complexities of the

Table 5.2. Looking at What a Second-Tour Leader Does

Army-Wide Behavior Scales:
 1. Demonstrating Technical Knowledge and Skill
 2. Demonstrating Effort
 3. Supervising Subordinates
 4. Following Regulations and Orders
 5. Demonstrating Integrity
 6. Training and Developing Subordinates
 7. Maintaining Equipment
 8. Physical Fitness
 9. Self-Development
 10. Showing Consideration for Subordinates
 11. Demonstrating Appropriate Military Bearing
 12. Demonstrating Appropriate Self-Control

Additional Leadership Scales:
 13. Serving as a Role Model
 14. Communication With Subordinates
 15. Personal Counseling
 16. Monitoring Subordinate Performance
 17. Organizing Missions/Operations
 18. Personnel Administration
 19. Performance Counseling

General Scales:
 20. Overall Effectiveness
 21. Senior NCO Potential

emerging Army. The axiom has it that the best predictor of future performance is past performance. True, yet the situation is not quite that simple. Records of past performance may not be that readily available -- or may not be that relevant if a reenlisting soldier wants to change MOS, which often happens. Furthermore, a more senior noncom will face demands to lead, manage, and teach that were not part of his or her experience. So one turns again to that first indicator of aptitude -- scores on the ASVAB at time of applying to the Army.

To see how closely ASVAB scores might be related to second-tour performance, in 1985 the Army matched the records of all soldiers who had entered active duty between 1 Jan 76 and 30 Sep 80 against their Skill Qualification Test level scores for 1983 and 1984. The search yielded 195 skill level samples from 95 MOS that were large enough for analysis of how these soldiers were faring as noncoms. The ASVAB selection (predictor) records used were the AFQT, the General Technical composite (the Verbal and Arithmetic reasoning subtests, commonly regarded as a measure of general aptitude), and the ASVAB Aptitude Area composites.[18]

Some 84 percent (163 of 195) of the samples showed significant positive relationships between GT predictor scores and the operational SQT scores; 96 percent (186 of 195) showed such relationships between the Aptitude Area scores (for their MOS) and SQT scores. Analyses indicated that use of an Aptitude Area cutoff score as one of the standards for reenlistment would be quite appropriate for most MOS.[18]

Sequential Prediction

A unique opportunity in personnel research -- to test a large number of job applicants and follow them from selection for three to five years -- through training, job performance, and promotion -- was afforded by the Project A/Career Force multi-year program. The controlled sequence of development, testing, and followup let military researchers look at what happens to performance over time and experience. How well do various predictor tests, given at the time of enlistment, match up with the soldier's success at various stages in the Army life cycle? How well does a soldier's performance in training predict later performance on the job? And how well does a soldier's first-tour level of performance predict behavior as a junior NCO in a second tour?

The plunging MOS sample sizes between accession and second-tour in Table 5.3 provide a rough indication of how hard it is to follow (and be able to test) one particular soldier from before Basic Training to after reenlistment. The total number tested at the end of training was over 26,000; the total from that number who were tested during their second tour was under 1,600. The shrinkage is testimony to attrition, ineligibility or failure to reenlist, and the sheer difficulty of trying to locate and meet with groups of those soldiers scattered all over the world.

Figures 5.1 and 5.2 illustrate the degree of success in predicting training and first-tour performance.[37] As noted earlier, ASVAB is excellent at predicting technical and general soldiering skills, quite good at predicting effort and achievement, and modest in predicting the other first-term performance dimensions. New tests in the Experimental Predictor Battery do not add very much to the ASVAB in predicting skills but add considerably in the other dimensions of performance.

Table 5.3. Sample Sizes for Batch A MOS at Four Points Over Six Years

MOS		Prediction	Performance[a]		
		Accession	End of Training	During First Tour	During Second Tour
11B	Infantryman	14,200	8,117	909	47
13B	Cannon Crewmember	5,100	4,712	916	80
19E/K	Armor Crewman	2,400	2,048	1,073	168
31C	Single Channel Radio Operator	1,100	667	529	60
63B	Light-Wheel Vehicle Mechanic	2,200	1,215	752	194
71L	Administrative Specialist	2,100	1,414	678	157
88M	Motor Transport Operator	1,600	1,354	682	89
91A/B	Medical Specialist	7,700	3,218	824	222
95B	Military Police	4,200	3,639	452	168
	Total	40,600	26,384	6,815	1,595

[a] Accession figures are summarized from data in the Project A Final Report[10]; performance test figures are from the Career Force Final Report[11].

Figure 5.1. Pre-Enlistment Prediction of Training Performance
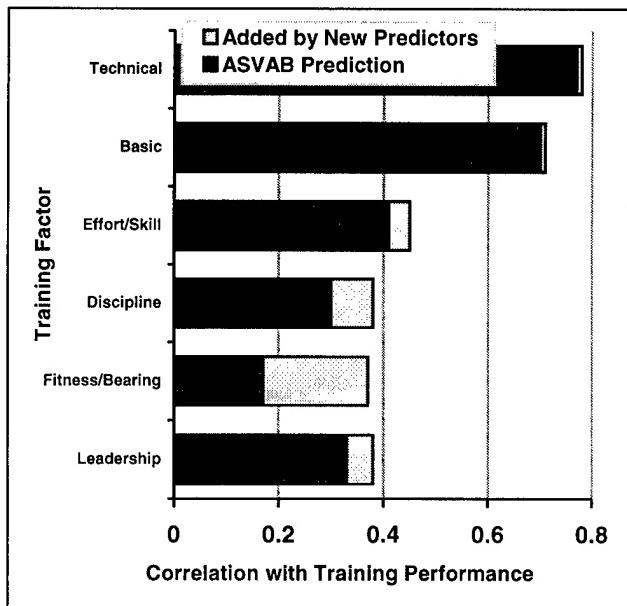
Figure 5.2. Pre-Enlistment Prediction of First-Tour Performance
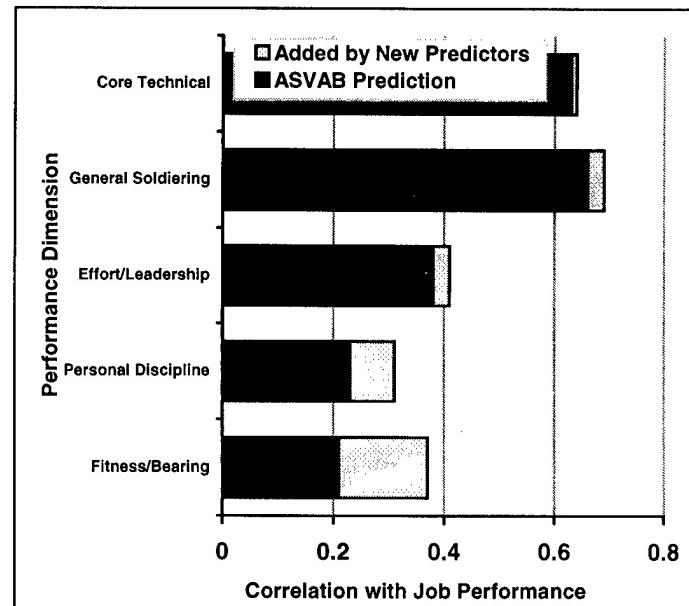
The relationships between performances assessed at different stages are substantial. Performance in training does predict performance on a first-tour job (Figure 5.3), and performance in the first tour of duty does predict performance in the second tour. For the second tour, the ASVAB pre-enlistment score is by far the strongest predictor of "Can-Do" performance. In contrast, adding the other roll-up measures to ASVAB more than doubles prediction of second tour "Will-Do" performance.

Figure 5.3. Training Performance Prediction of First-Tour Performance
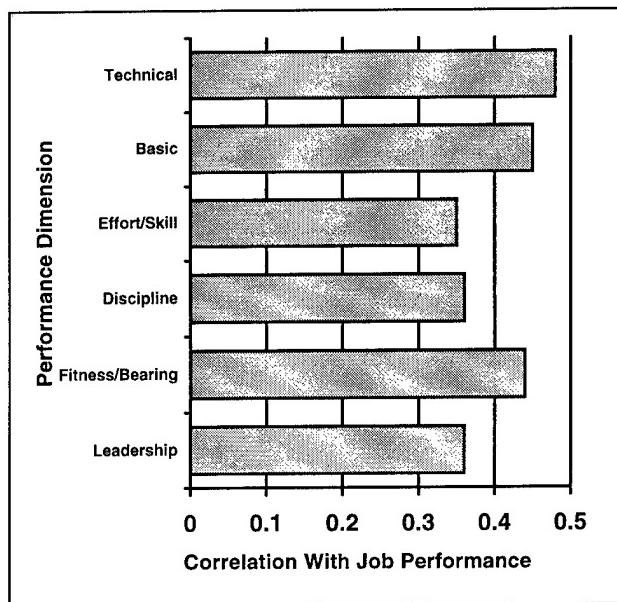
It is important that NCO potential -- the Leadership element -- is predicted by every kind of predictor. The same thing happens when analysts try to map the traits that constitute leadership -- everything seems to contribute. Leadership seems to be a combination of many attributes; so the selection and training of leaders must take a full set of capabilities into account.[11]

Figure 5.4. "Roll-Up" Prediction of Performance
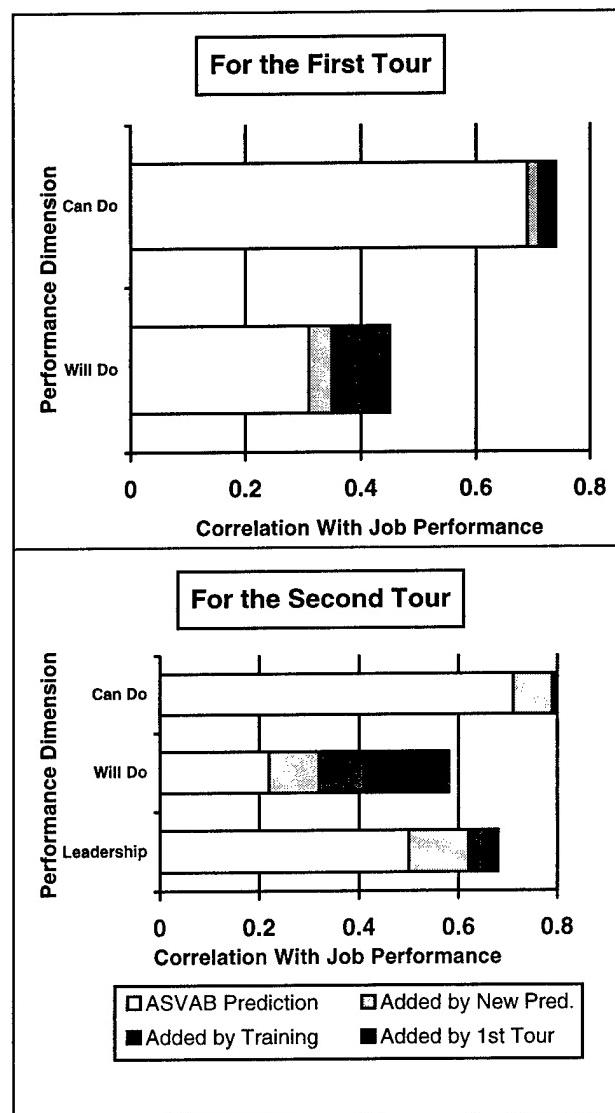
The phases of the "roll-up" analyses are pulled together in Figure 5.4 for the first tour and the second tour. The figure indicates how much predictive strength is added by each stage of performance measurement.

Rolling up" the earlier information into later, richer predictor sets results in progressively more accurate prediction of future performance.[27] Some conclusions from the "roll-up":

- When predicting "Can-Do" criteria, it is difficult to improve upon the initial ASVAB and Experimental Battery predictions, although training data help predict first-tour performance.

- Both the Experimental Battery measures and prior performance contribute substantially to predicting second-tour performance.

- In general, the "roll-up" values are quite high, even when predicting NCO second-tour leadership from test battery data obtained some six years earlier.

There were many lessons learned in the early years of the all-volunteer Army, but generally they can be distilled into a few key principles:

- Recruiting and retaining quality soldiers pays off on the battlefield.

- The up-front cost of recruiting quality soldiers seems high, but the long-term costs of lowering quality are much, much higher.

MG Kenneth W. Simpson
Commanding General,
U.S. Army Recruiting Command

This
Is
The
Army!
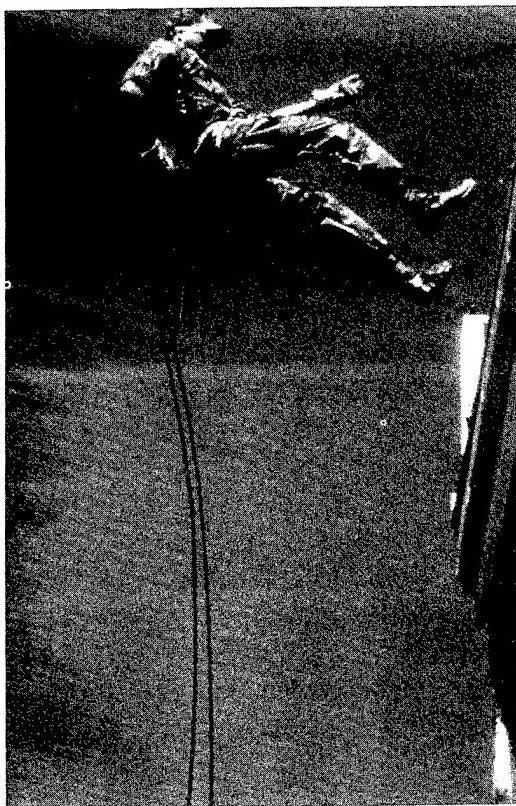
Special Topics: Attrition, Low Aptitude, and Combat Performance

6

Up to now, we have examined the nature of first- and second-tour performance and seen how various human qualities underlie success in Army jobs. These findings from Project A/Career Force are complemented by research on three other perennial concerns in the Army's personnel system:

- First, knowing that ASVAB predicts performance out into the second tour, how can we select soldiers to maximize the odds that they will complete the first tour? This is the problem of attrition.

- Second, in times of mobilization, the Army takes in large numbers of citizens in the lower ranges of general aptitude. How can we select and assign them most effectively?

- And finally, all these research results are very encouraging, but what do they tell us about ultimate Army job performance -- in combat?

The Anatomy of Attrition

The effort to predict successful performance in the Army has a shadow: How to keep from choosing applicants who will not succeed. First-term attrition -- failure of an enlistee to complete his or her obligated term of service -- incurs great costs for the Army. It is also highly visible; from one-fourth to one-third of enlistees do not finish that first tour (Table 6.1).

**Table 6.1. Attrition Rates
(Project A Longitudinal Sample)**

Variable	Number	Attrition Rate (Percent)
Age		
High School Graduates		
17	1,708	22
18-20	24,430	23
21-25	7,387	27
26 +	1,976	33
Non-Graduates		
17	597	54
18-20	1,630	47
21-25	466	42
26 +	120	39
AFQT Category		
I-IIIA	23,991	25
IIIB	10,561	28
IV	3,762	27
High School Degree		
Diploma	35,501	24
GED	857	43
None	1,956	49
Enlistment Term		
2 years	4,881	12
3-4 years	33,433	28
Gender		
Female	3,824	34
Male	34,490	25
Race		
White	27,749	27
Black	7,904	24
Hispanic	1,445	20
Other	1,216	19

Some attrition of course has to be viewed as unpredictable and unavoidable because of things outside the control of the soldier or

---

Among the driving forces for [adaptability screening] are the high monetary costs, decrease in readiness, and general turbulence created when many of the forces fail to adapt and are separated from the service prior to completing an obligated term of enlistment. New means of controlling personnel attrition have been investigated with the hopes of quelling . . . adaptability problems. . . .

The costs mount up because of lost investment in training, higher recruiting and salary costs, veterans' benefits expenditures, and payment of unemployment compensation to separated Service members. There are also non-pecuniary or indirect costs which include force instability, lowered morale, and lack of readiness. Individuals may also pay a price. . . Failure in military service may significantly affect their future employment opportunities and earning potential.

Education Standards and Military Selection[25]

---

the Army, such as unforeseen medical or family problems. What is of concern is the proportion of attrition for reasons like low satisfaction (on either side), disciplinary problems, inept performance -- the kind of happenings that might have been foreseen from screening information about an applicant's interests and previous behavior.

More than 30 years ago, an inquiry into the causes of attrition revealed that level of education stood out as the single best predictor of military adjustment.[17] Similar findings surfaced repeatedly during the intervening years, and continue today: High school diploma graduates are considerably more

likely than non-graduates to complete their first-term military contract.

In contrast, the relationship between attrition and measures of cognitive ability, such as AFQT and ASVAB, does not show marked differences between aptitude levels. In fact, there have been times when soldiers who were in the lowest acceptable portion of the AFQT (Category IV) but who were high school graduates have had lower attrition rates than the topmost level in the AFQT rankings. One military analyst[25] muses:

If it's not academic ability that separates the credential holders, what are the attributes that account for whether a recruit is likely to be a stayer or a leaver? . . . Receipt of a traditional diploma could signify perseverance, maturity, participation in and adaptability to rules and regulations, determination, self-control, and other similar attributes. Finishing high school may socialize one appropriately for aspects of the military environment. After all, sitting in a classroom is part of the training process and rules and regulations are a fact of service life. . . . Teamwork is of the essence in the military, and conformity and self-control are drilled. Whether and which of these qualities are involved in the diploma's predictiveness [is] difficult to measure, let alone affirm their relationship to attrition.

These are the kinds of characteristics that are explored in temperament/biodata/interest measures like ABLE and AVOICE. How well such instruments predict attrition, in addition to their prediction of performance, had a substantial place on the Project A/Career Force agenda. Some of the findings relating temperament scores with AFQT category and level of education are shown in Figure 6.1.[53]

When and why recruits separated from the Army before the end of their first tour was studied in Career Force[28] for four job groups (combat/noncombat MOS, 3- or 4-year enlistment terms). For example, Figure 6.2

Figure 6.1. Predicted Attrition, by Temperament Score
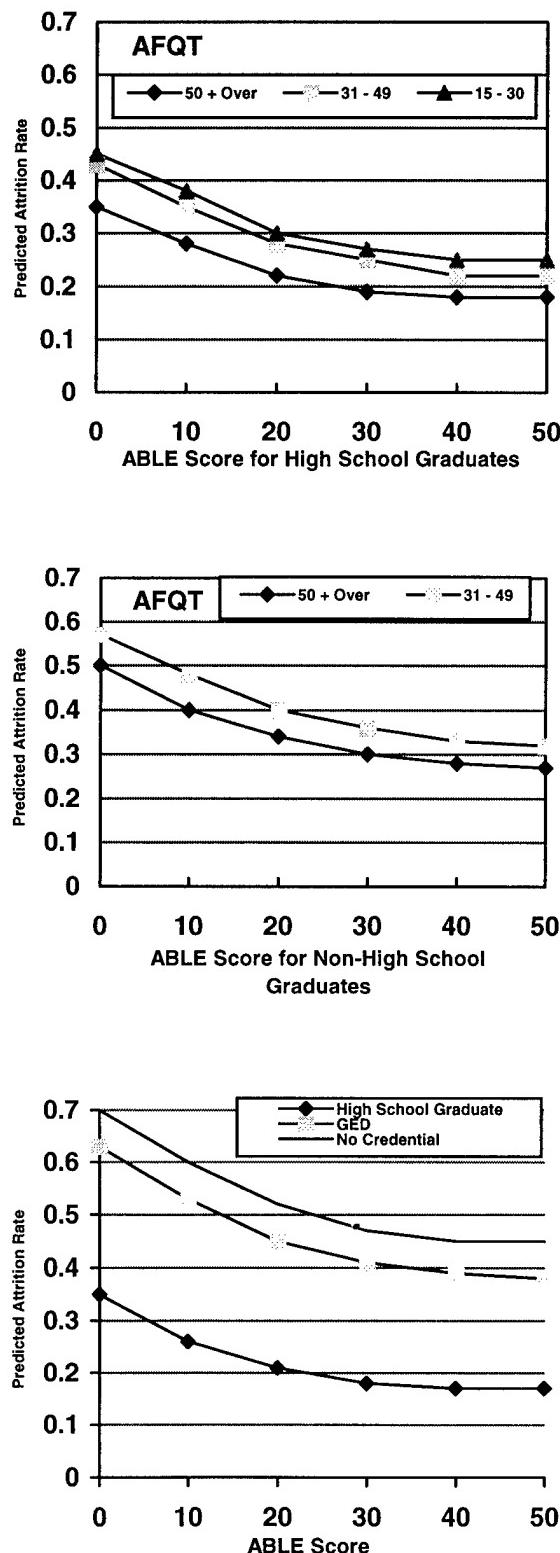
(data) from one of the MOS) shows that 20 percent of the non-graduates had been lost to the Army in 10 months, whereas high school graduates did not reach this loss figure until 24 months. By the end of the 3-year tour, twice as many non-graduates as graduates had failed to finish.

To see what the experimental predictors might contribute to lessening attrition over the full first tour, predictive data for the four job groups were analyzed. The base predictor group--the four ASVAB factors and high school diploma graduate status -- proved to be significantly related to first-term attrition. Addition of ABLE temperament scores provided a significant increase in prediction for all job groups. AVOICE interests measures increased prediction for 3-year but not for 4-year enlistments.[28]

Trying out various combinations of tests, the best subset across all four job groups was composed of high school diploma graduate status and five ABLE scales. Within the set of scales, Emotional Stability was the best predictor of one-year attrition, while Nondelinquency was the best predictor of 3-year attrition.[53] The Quantitative composite (math-related) from the ASVAB occurred in three of the four top subsets. (See Figure 6.3[40]) These seven variables can be seen as a single composite for predicting first-term Army attrition.

The overall results reaffirm that high school diploma graduate status is a powerful predictor, with the attrition rate for non-graduates being approximately twice that of graduates. Other factors are involved, however, and the kind of information provided by ABLE measures of temperament can add significantly in identifying the attrition-prone before enlistment.

Figure 6.2 Continuation Rates for Graduates and Nongraduates (MOS 11B)
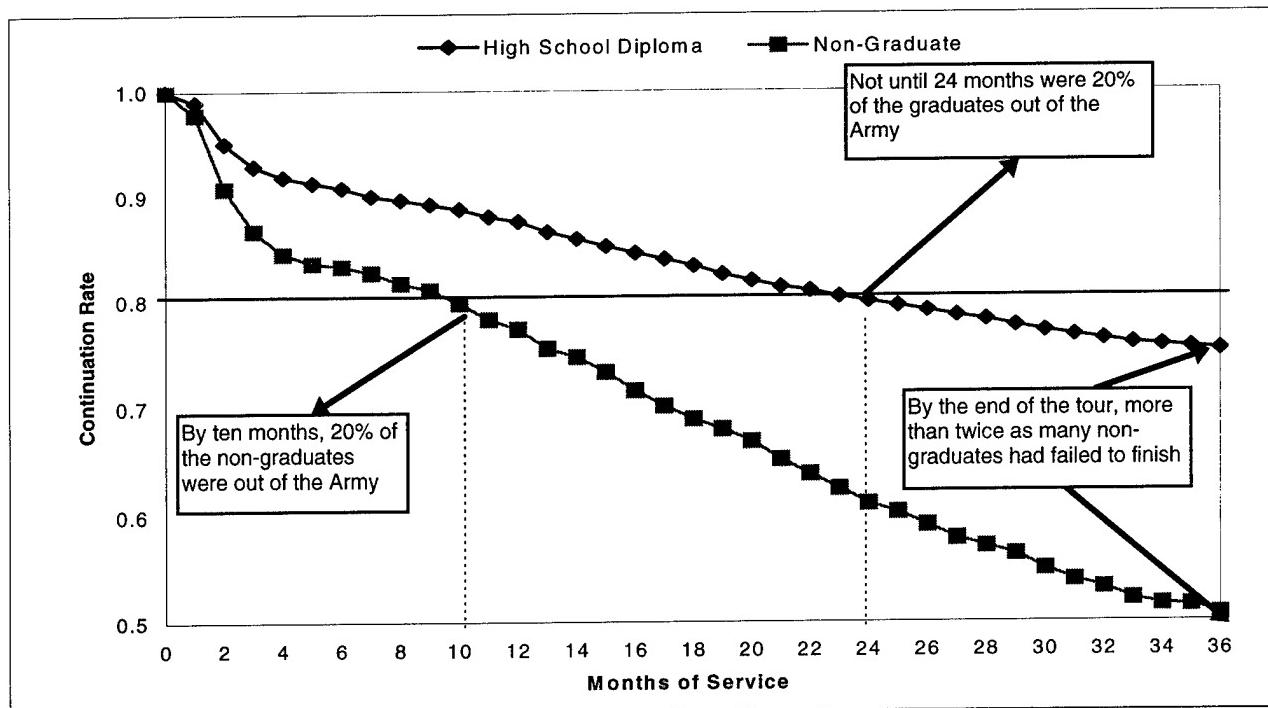
Figure 6.3. Best Subset Predictors of Attrition Through First Tour

- Strongest Predictor: High School Graduate Status

- Significant Contributions from these ABLE Temperament Scales:
  - Nondelinquency
  - Dominance
  - Physical Condition
  - Self-Esteem
  - Social Desirability

- Some Contribution from ASVAB Quantitative Composite:
  - For 3-Year Enlistments - Significant
  - For 4-Year Enlistments - Significant for Combat MOS Only

Another approach to the attrition problem was an effort to find out how -- and how much -- soldiers' satisfaction with their jobs might be related to performance, attrition, and turnover (defined here as failure to reenlist although eligible). The Army Job Satisfaction Questionnaire (AJSQ) was administered to Project A/Career Force soldiers during their first tour. Making use of experience with earlier military materials on this topic, the AJSQ covered six aspects of job satisfaction: with supervisors, co-workers, promotions, pay, work, and the Army as an organization. It also dealt with reasons for enlistment and intentions about reenlistment.

Because this survey was given to soldiers after 12 to 24 months of service, it did not reveal reasons for earlier attrition. But it did yield background information on how soldiers felt about various aspects of their work which were found to be predictive of late-term attrition.[23]

Utilizing Low-Aptitude Soldiers

The shifting composition and declining size of the youth manpower pool have placed a premium on looking closely at how low-aptitude personnel have performed in the Army. There have been two main periods in recent decades during which large numbers of persons in AFQT Category IV (low-aptitude personnel) served in the Army.

The first was the Project 100,000 experiment, begun in 1966 as part of the War on Poverty. The intent was to aid the disadvantaged through military service while evening out the benefits and burdens of such service, and to get experience with low-aptitude soldiers for policy planning. However, the Vietnam War and the variety of ways in which the so-called "New Standards Men" were trained and placed confused the

picture, and interest in the project declined as the war demands faded.[35]

The second episode was the inadvertent admission of some 300,000 low-aptitude recruits during the misnorming of the ASVAB (1976-1980), which inflated AFQT scores in the below-average range. Their in-service records, compared to higher aptitude soldiers, showed higher basic training dropout rates, lower Skill Qualification Test scores, lower promotion rates, higher attrition, and higher disciplinary levels. Nevertheless, a significant proportion of soldiers in Category IV were able to reach an adequate level of performance in a reasonable time, and 14 percent of them were still on active duty at the end of 1988.[35]

To develop specific "lessons learned" from this experience with low-aptitude men, the Army set up several programs aimed at isolating individual traits that might help predict job success for those of lesser aptitude, and identifying MOS in which soldiers of average and below-average ability can perform most and least effectively.[29]

As performance measures for these purposes, the project used "end-results" -- attrition, SQT score, promotion rate, and reenlistment eligibility -- drawing data from files of the 1977-80 cohort. Twenty-five MOS considered to be representative were selected from the 260-plus entry-level MOS at that time. A wide variety of predictors (ASVAB and others) was chosen to sort out individual characteristics.[29]

The performance record of the 1977-80 accessions in all aptitude categories is shown in Table 6.2. As in the earlier analysis, soldiers in AFQT Category IV, on the average, had lower job knowledge scores. They were more likely to leave prematurely, and were less likely to be promoted or eligible to

Table 6.2. Performance of 1977-1980 Army Accessions by AFQT Category (25 MOS)

Measure	AFQT Category				
	I & II	IIIA	IIIB	IVA-IVB	IVC
Attrition (%)	26.8	34.3	38.4	39.9	28.4
Promotion E-4 (%)	67.0	59.9	56.3	54.7	65.7
Promotion E-5 (%)	15.3	9.0	7.2	6.1	4.3
SQT Scores	112.8	105.8	102.0	98.0	93.6
Reenlistment (%)	27.5	27.0	26.1	26.4	32.1

reenlist, but substantial numbers performed adequately. The fact that only soldiers in Category IVC were required to be high school graduates probably accounts for their surpassing soldiers in IVA-IVB in some areas.[29]

The many detailed predictor-criterion comparisons for the AFQT categories and the 25 MOS do not lend themselves to generalizations but provide much to consider about placement. Again, high school graduation status was the best single predictor. For lower aptitude soldiers the vocationally oriented ASVAB subtests such as Mechanical Comprehension tended to be stronger predictors than the verbal and math subtests in the AFQT. Certain types of low-complexity MOS emerged as more suitable for lower-aptitude soldiers. Some of the predictors measuring temperament appeared to be especially informative for individuals of low aptitude.[29]

All in all, the information gathered provides a good starting point for a selection that allows traits of temperament to compensate somewhat for lower AFQT.[35] Such a process would allow low-aptitude applicants to enlist if they met additional standards that are not necessarily required of high-aptitude applicants.

Performance in Combat

Preparation for combat is the basic theme and combat readiness is the basic goal for all Army activities during peacetime. Vast amounts of information are available on these matters. But how is soldiers' performance in peacetime related to actual performance in combat? That's not so easily answered.

In fact, the last systematic study of relationships between combat and peacetime performance of American troops was conducted 40 years ago during the Korean War.[54] In those studies, more than 300 soldiers in Korea who had been classed as effective or ineffective on the basis of recent behavior in combat were interviewed and tested on a wide variety of personality, vocational interests, and ability measures; results showed, for example, that effective fighters were more emotionally stable, dominant, and socially mature.[15]

In contrast, there is extensive research on the structure and qualities of peacetime performance. Selection and classification measures used during enlistment screening are proven predictors of job performance in peacetime. However, their validity against combat criteria had not been well established.[54]

Why is AFQT score considered so important? In part, because AFQT score is believed to represent cognitive abilities (e.g., reasoning and learning). The Army believes that soldiers with high AFQT scores are likely to be especially capable in dealing with unfamiliar or stressful situations, where they must quickly assimilate new information and make decisions rapidly. Such unexpected situations, of course, are frequently encountered in combat. High-scoring soldiers are also expected to be better able to perform complex tasks demanded by modern military equipment, which is increasingly entering the U.S. inventory.

A considerable body of empirical research has documented that high-AFQT soldiers do better at a wide range of specific tasks. ... Congress, however, has been interested in obtaining more direct measures of AFQT effects on unit readiness and warfighting capability. To date, such evidence has been scattered and to some extent anecdotal.

Soldier Performance Research Project[3]

Thus an extraordinary opportunity came about when the Army units deployed in Operation Desert Shield/Storm in 1990 included soldiers who had been part of the Project A sample and had reenlisted; their enlistment predictor scores and their first-tour performance records were therefore available. Ratings of their performance in combat were obtained in 1992 when Career Force second-tour measurements were being administered to these soldiers. A total of 204 soldiers, 57 percent of them from combat MOS, were rated by supervisors who were familiar with their combat performance.

The 25 items in the ratings were formed into four composite outcomes -- Leadership in Battle, Task Proficiency, Supervision, and Personal Discipline -- plus an item on handling physical demands and an overall rating of effectiveness in Desert Shield/Storm.

Some of the findings:[54]

Relationship Between Peacetime and Combat Performance. Soldiers' level of Core Technical Proficiency in first tour was correlated with skilled task performance during Desert Shield/Storm, as were Effort/Leadership and Personal Discipline. Soldiers who had fewer disciplinary incidents in the garrison showed higher levels of personal discipline in combat. Relationships between performance in combat and after the war showed the same pattern; the strongest connections were between combat performance and Effort/Achievement and Personal Discipline.

ASVAB and Combat Performance. The ASVAB scores at enlistment some five years earlier were significantly correlated with task proficiency and supervision during Desert Shield/Storm. As in peacetime, ASVAB was not related to ratings of personal discipline or physical performance. However, scores on the ASVAB Speed factor proved to be significantly related to leadership effectiveness in battle.

Temperament and Combat Performance. Relationships between the ABLE measures administered at time of entry and again after combat experience were low. In contrast, soldiers' Dominance, Achievement, and Adjustment assessed closer to the conflict were significantly related to combat performance.

In summary, strong relationships were evident between performance in combat and performance measured before and after the war. The ASVAB, at enlistment, was predictive of task proficiency and supervision during combat. And, soldiers' levels of Achievement and Dominance, assessed after

the war, were positively related to combat performance.

> The challenge we face is to determine what kind of research will enable us to select the persons who will be able to act and lead effectively in the emerging uncertain world.
>
> MG Fred Gorden
> Director of Military Personnel Management

As a surrogate for combat performance data, combat simulation and combat-related activities are used for training and research to further the goal of combat readiness. One program of this type was the Soldier Performance Research Project (SPRP), which was a response to a Congressional inquiry about the relation of AFQT levels to combat performance and readiness.

M1 tank gunnery was the subject of one SPRP project.[3] In a test of 547 trainees at Fort Knox, soldiers in AFQT Categories I and II hit 67 percent of the targets while those in Category IV hit only 53 percent; also, Category IV soldiers were significantly slower on opening times and lower in hits per minute. In a test of M1 tactical operations, results were similar. Summary findings from these and other SPRP studies are shown in Figure 6.4; all show a similar decline in performance with declining AFQT. The potency of ASVAB shows again in this combat-like performance.

Figure 6.4. Some Findings From the Soldier Performance Research Project

M1 Tank Gunnery

	Mental Category			
	I&II	IIIA	IIIB	IV
Percent Hits	67.1	64.1	59.1	53.3
Opening Time (Seconds)	17.7	18.3	18.9	19.7
Speed/Accuracy	53.6	50.0	47.6	43.9

M1 Tank Tactical Operations

	I&II	IIIA	IIIB	IV
Tank Commander	.53	.52	.46	.43
Driver	.52	.47	.50	.45
Relative to I&II TCs (%)	100	98	87	81

TOW Anti-Tank Gunnery
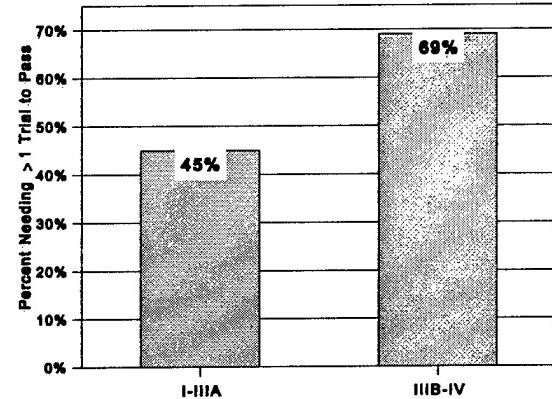
M1 Tank Field Test
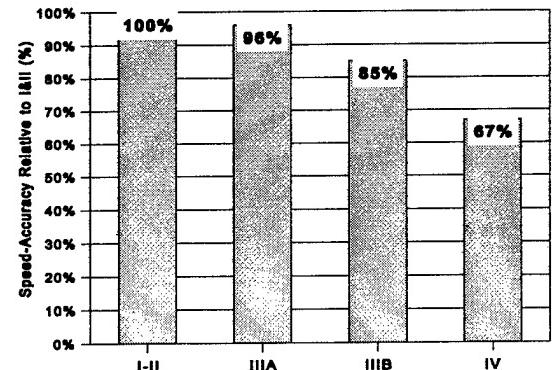
M1 Tank Simulated Networking Test

	I&II	IIIA	IIIB	IV
Tank Commander	.61	.59	.58	.53
Driver	.62	.56	.57	.56
Relative to I&II TCs (%)	100	97	95	87

Improving Soldier Selection

It is one thing to get solid research results under controlled conditions. It is quite another thing to put the findings to work in an operational military world where the people involved number in the hundreds of thousands.

The complexities in implementing research products are formidable enough if it involves only the Army. But if the potential actions would affect all the Services, as is true for possible changes in the ASVAB, these complexities multiply. Furthermore, any change involves a lot of work and may be expensive. Those facts of life provide background for the discussions that follow.

Possible Additions to the ASVAB

Much of the selection and classification research of recent years grew out of that basic inquiry: How good is the ASVAB? Answers went in two directions:

What does the ASVAB measure and how well does it do this?

The answer: The ASVAB measures cognitive aptitudes (verbal, quantitative, technical, and speed). Individuals' scores on ASVAB before enlistment very strongly predict their technical performance and leadership in the Army. ASVAB's power to predict personal discipline, fitness, and military bearing is moderate. In contrast, scores on ASVAB are not consistently related to first-tour attrition.

What does the ASVAB not measure?

The answer: It does not measure spatial and psychomotor capabilities that might be very relevant for a soldier. It does not assess personality, life experiences, and personal interests that might provide strong clues to the individual's prospects.

The problem in selection and classification is how best to assign available personnel resources to meet the near term demands on the personnel system. The problem in force modernization is to predict if the long term personnel demands imposed by improved and new weapons systems requirements will exceed the available manpower and personnel resources.
The Army's Personnel Problem . . . [45]

The goal of supported research within the military [is] to bring about cost-effective change within the Army.
Behavioral Science in the Army [58]

Potential Benefits of New Measures

By adding new types of tests to the ASVAB, the Army could realize the benefits of higher job proficiency through a better fit between applicants and military jobs, and substantial savings in time and money through lower attrition by "selecting out" the unsuitable. But the benefits and pitfalls of implementing new tests vary with the type of measure.

Spatial and Psychomotor Measures. Research summarized in Chapters 3-6 indicated that testing such attributes as spatial orientation and perceptual precision brought small but consistent gains over ASVAB alone in predicting technical proficiency -- the "Can-Do" performance factors. Such gains turn out to be very valuable in the Army, multiplied as they are by the tens of thousands of incoming recruits. The psychomotor tests have already

proven to be capable of markedly improving accuracy in tank and anti-tank gunnery. As a result of DoD's joint service evaluation of new computerized tests, some of the Project A spatial tests, now computerized, have been selected for further tryout as future ASVAB subtests. The practicality of adding new tests to ASVAB has improved now that DoD is converting to a faster, computerized version of it.

Temperament, Biodata, and Interests. Personal information from the ABLE and AVOICE questionnaires added substantially to ASVAB in predicting attrition and first-tour success in personal discipline and physical fitness -- the "Will-Do" performance factors. These predictions have little relationship to ASVAB scores and education level, so they add new power for identifying successful future first-tour soldiers. Versions of the ABLE have already been used with success in small-scale applications.

An issue that bears on the feasibility of this type of self-reported measures in large-scale testing is the possibility that examinees will try to make themselves look too good (or too bad) when answering. Considerable research has been done on ways to counteract or control such faking.[37] Results have indicated that respondents can slant their answers to make themselves look unrealistically good or bad. However, special scales in ABLE, designed to detect distorted or inaccurate answers, are quite successful in identifying examinees who are less than candid.[20] Further work is underway on how to apply these scales and on developing questions that are less vulnerable to faking.

Balancing Effectiveness Against Cost of Change
Would change be worth it? That's where decision time comes for the plans-and-policies

levels of the military command structure. One can get a feel for the magnitude of that decision by looking at the numbers and dollars for one recent Army cohort, from the recruiter's office through reenlistment, shown in Figure 7.1.

Cost-benefit analyses of selection success are complex and arguable even in civilian business settings, and the military environment adds a whole new range of perplexities. The most serious: How can a dollar value be put on successful selection when the impact is possible life-or-death for the individual and victory-or-defeat for the nation? Yet there does have to be some way of assessing the usefulness of performance, at different levels and in different jobs. Choices do have to be made.

Efforts to find ways of putting a value on military performance have used various approaches.[14,42] The Selection and Classification Models Project developed and compared alternative methods of evaluation. A cost-effectiveness framework was set up for estimating the value of selection and classification testing in terms of the dollar costs to recruit, train, and pay a first-term enlisted force operating at various levels of expected performance.[19]

This framework used performance prediction equations for nine occupational areas computed from a given test battery. These were combined with recruiting, training, and compensation costs and analyzed for the most cost-effective mix that met the performance goals for each job family. Data from the Project A database were used to evaluate four predictor test batteries: A--the AFQT (math/verbal tests); B--the four ASVAB factors (Quantitative, Technical, Verbal, and Speed); C--a spatial/psychomotor

Figure 7.1. Cohort Numbers (FY93) and Dollars: Army Enlisted Personnel Flow

Appointments	Tested	Accessed	Cost	
			Incremental	Total

(Figure content:)

Appointments 460,000 → Tested 144,000 → Accessed 78,000 — Incremental $590M, Total $590M

Training 70,200 — Incremental $1,490M, Total $2,080M

Complete Term 49,920 → Reenlist 19,030 — Incremental $4,450M, Total $6,530M

= 19,030

August 1994
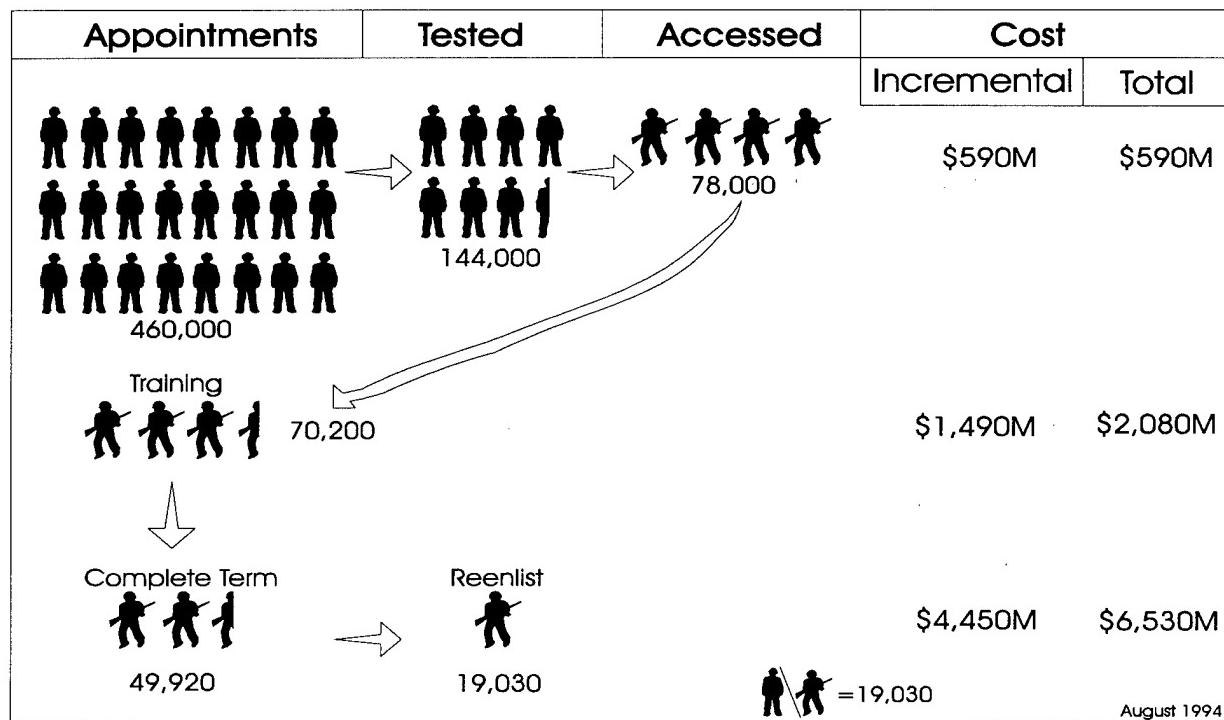
composite added to the ASVAB; D--the ABLE temperament scales added to battery C. The potential value of improved testing was estimated as the reduction in total cost necessary to meet the established performance goals.[19]

The resulting estimates suggested that adding a spatial measure to the ASVAB may save up to $114 million in recruiting, training, and compensation costs for an Army cohort over four years (Figure 7.2). The results also indicated, again, that spatial aptitude would be particularly useful in finding occupational areas where lower AFQT recruits would function well. Adding ABLE's temperament measure to battery C was estimated to save an additional $159 million.[19]

This evaluation framework can be applied to a number of different policy issues facing the Army. For example: How would results change if realistic factors such as training seat availability were added to the simulations? What are the expected costs of eliminating a test, such as Numerical Operations, from the current selection battery? What is the benefit (and cost) of giving applicants more (or fewer) Army jobs to choose from?[19]

Questions About "Quality"

What "quality" mix does the Army need? The response to that question may depend on when it is asked. In wartime, the need to bring in very large numbers may require lowering of aptitude standards. Currently the rapid growth of technology is making the Army eager to enlist more high-aptitude youth. The civilian job market has an impact: when high aptitude youth have plentiful civilian job opportunities, the Army will take more lower aptitude applicants and make up the difference with training and leadership. And if the youth manpower pool itself changes, with a higher proportion of young people having limited language skills and

Figure 7.2. Estimated Annual Benefits From New Selection Tests

- **Basis for Estimating Benefits**

  - Determine Total Costs (Recruiting + Training + Compensation) to Reach a Particular Performance Level

  - Determine Degree of Improvement Generated from New Tests

  - Determine Costs Required to Reach Same Level of Improvement

- **Estimated Benefits (Savings) Associated With New Tests**

  - Psychomotor/Spatial tests: $114 Million

  - ABLE: $159 Million

  - Total: $273 Million

Note: Benefits based on results from Project A concurrent validation.[37]

poorer education, the Army will have to find ways to cope with the available supply of youth.

All that aside, having high-aptitude enlistees naturally has many advantages for training, flexibility of assignment, and general career promise. It is clear that, on average, higher aptitude soldiers outperform soldiers of lower aptitude. Still, recruiting costs are higher for high-aptitude prospects and the competition is heavy. Many military jobs can be -- and are -- performed very competently by soldiers of mid- and lower aptitude. It is very much to the point that Army aptitude refers not to intelligence as such but to trainability. Also, it is important to select applicants in a way that reduces costly attrition; but aptitude scores are not particularly related to attrition. So the question of the aptitude mix that is desirable has more than one answer.

The Services are moving away from highly specialized jobs in their downsizing by consolidating MOS.[9] Such a trend has many implications -- some of them conflicting -- for training, the setting of MOS standards, and future aptitude requirements. As MOS are combined, enlistees will need to be more versatile and capable of handling diverse tasks.[9] In such a situation, measures of "learning rate" might predict performance in the new generalist jobs.[24]

Continuing improvements in MOS classification systems, discussed below, are closely linked to making better use of soldiers at all aptitude levels. Fuller identification of the detailed and specialized factors in performance -- a major element in much of the research that has been described here -- provides the kind of information needed to select and assign soldiers in accordance with their individual abilities.

Matching Soldiers to Jobs

One of the major benefits from this personnel research program has been the emergence of both concepts and techniques that can tell us more about the strengths and weaknesses of the young person who is contemplating joining the Army. Such a capability greatly increases the prospect not only of more informed selection but also of a job-person match that makes the most of each individual's potential. Over time, the effect will be both higher efficiency and lower attrition for the Army.

The Second Stage of Selection: Classification

Look at two different people doing two different jobs. On one job, A succeeds and B fails; on the other, A fails and B succeeds. The payoff in classification is getting A assigned to the first job and B assigned to the second job. Our increasing ability to take individual differences into account in making assignments leads many military planners to believe that major improvements are possible in classifying Army entry-level soldiers.

It needs to be recognized that the "classification problem" is almost uniquely military. In civilian organizations, it is much more typical for applicants to be considered for only one, or perhaps two jobs. Having more than 200 possible job assignments for a single Army applicant is a vastly different situation. Moreover, the research on optimal classification has been very limited because such research does not have much commercial application and because it requires complex new computer simulations and computations.

Research and development efforts on classification at ARI have taken two complementary directions in recent years. The first of these is a more powerful software system, the Enlisted Personnel Allocation

System (EPAS).[23] It uses optimizing routines to increase the Army's flexibility in assigning individuals to MOS (by way of training seats) through three innovations: extending the time horizon over which the matching is done; taking into account both the projected and the known supply of applicants; and making job assignments based on individuals' *highest* Aptitude Area Composite scores rather than minimums.

The second approach focuses on how MOS are grouped for classification and on how the Aptitude Area Composite scores themselves (i.e., of the *battery* of tests) could be improved. This research has suggested that a 200 percent improvement in total Army job performance might be gained from better use of the current ASVAB subtests and that even greater improvement would be realized by supplementing ASVAB with new tests from Project A.[21,43,59,60]

According to an Army Science Board report by Alexander (1980),[2] it is not enough for a research community to exist, or even for it to be working on problems of concern to policymakers. At some level, applied research must be directly tied to specific questions for which knowledge of alternative courses of action and their consequences would be very valuable . . . . a special type of researcher is required, one who understands both the research and the policy worlds. Translation of research to action also requires people in the organization who are sensitive to the analytical approach and to the potential contributions of research to policymaking.

Designing, Planning, and Selling Project A[46]

Setting MOS Standards

Closely linked to the use of the ASVAB Aptitude Areas in steering Army prospects into career fields are the standards set for admission into each MOS. These standards specify the cutoff scores for applicants. The raising or lowering of the cutoff score is the

mechanism for determining the minimum aptitude level of soldiers entering that MOS. Proponents often ask to have a cutoff score raised so they'll get a higher proportion of bright students; but this is not a cost-free change of a paper requirement. Such a change would have impacts throughout the system, such as raising recruiting requirements or depriving other MOS of top-quality students.

The effects of various cutoff points in an MOS were part of a Career Force inquiry linking specified pre-enlistment predictors with attrition during the course of a first-tour enlistment. The effects of shifting cutoff scores is illustrated in Figure 7.3. It shows the attrition rate for an actual Career Force sample of soldiers in combat MOS, and what the attrition rate would have been if, at enlistment time, the Army had refused the lowest 10 percent, lowest 25 percent, and lowest 33 percent of those who were actually accepted -- that is, if the Army had raised the cutoff score.[28] As it was, the number remaining had dropped to 80 percent in about 20 months, and to just over 70 percent at the end of the 36-month tour. If the MOS had had a higher cutoff score that rejected the lowest one-third of those who were actually accepted, the number of those remaining would not have dropped to 80 percent until the 30th month, and would have gone only a little lower by the end of the enlistment.

To set sensible cutoff scores, it is necessary to understand what the particular MOS requires in the way of job performance. What level of performance is not acceptable? This question must be answered to determine where to set the minimum cutoff score on the classification test. Several approaches to determining minimally acceptable levels of job performance were explored in the Army Synthetic Validation Project.[55,56] They found that different approaches can lead to very

different conclusions and that it is essential for the Army subject matter experts who provide the judgments to clearly understand the nature and implications of their task.

Screening for Low-Aptitude Applicants

How to make appropriate use of applicants in the lower aptitude categories has been a matter of interest to the Army ever since the periods during Project 100,000 and the ASVAB misnorming when large numbers of low-aptitude people served in the Army. Despite problems, substantial numbers of these lower aptitude people served competently. The Army not only must use current personnel effectively but may need to accept more applicants from lower aptitude levels if the youth manpower pool changes or if world political conditions require larger U.S. forces.

The Augmented Selection Criteria for Enlisted Personnel[29] project sought to identify the characteristics of low-aptitude men who can or cannot perform successfully in the Army. Over a wide range of predictors, demographic variables, and criterion measures, high school diploma was useful across all outcomes; AFQT was often useful, but the best combination of predictors for low-aptitude people varied across MOS and across outcomes (e.g., attrition, SQT scores, promotions). The answer is not simple.

Taking a different tack, focusing on the characteristics of the job rather than those of the soldier, another project tried to identify those MOS in which low-aptitude soldiers were likely to make the greatest contribution to the Army. Results suggested that occupations with routinized performance requirements, long training times, and low levels of complexity were most suitable for soldiers with below-average AFQT scores.[29]

Figure 7.3. Effects on Attrition of Raising the MOS Cutoff Score
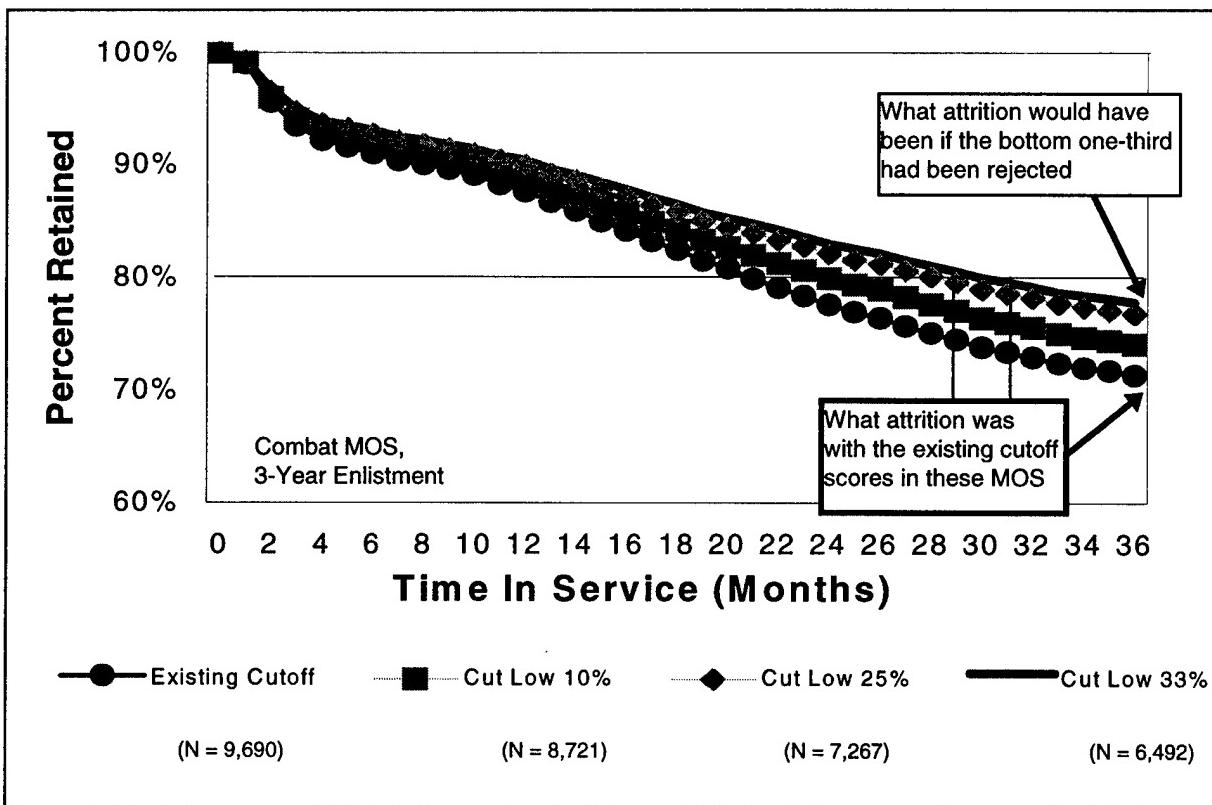
Conclusions

We have seen that the best use of a new predictor test depends on a variety of factors: which human attributes the test measures; which aspects of performance it predicts; what the available supply of youth is in numbers and aptitudes; and what types and levels of performance are required in the emerging Army of the 21st Century. Such factors feed into judgments of the costs and benefits of implementing the test.

A test for selecting among applicants would increase recruiting costs if it were used in a way that screened out higher proportions of youth. On the other hand, the same test could reduce such costs if it were used to

expand the set of human attributes that would qualify youth for enlistment.

For low impact on recruiting, a new test could be used to match qualified volunteers at all levels of aptitude with specific MOS; that is, for classification. Major gains in Army job performance can be expected from improvements in person-job matching through use of the new tests from Project A/Career Force along with new classification software developed by ARI.

These benefits depend, of course, on the tests being installed, used, and maintained. Bringing about that result has an added complexity: The Army does not have an agency with the responsibility for taking a

validated manpower technology (say, a new pre-enlistment test), carrying out further development to make the test usable (e.g., to conduct large sample data collections for developing norms and cut scores), developing any needed support (e.g., test administrator's resource materials), preparing the whole package for installation (e.g., getting central personnel databases changed to store the test's scores), and finally maintaining and renewing it (e.g., repeating the process periodically to produce fresh test forms). Such agencies exist to receive research products from the engineering counterparts of ARI and to transition those products into use. These agencies, the R&D Engineering Centers, are a link in the Army Life Cycle System Management Model (LCSMM) that is missing for the products of personnel-related research.

To realize the gains in Army job performance that Project A/Career Force has brought within the Army's reach, ARI is working with sponsors and users to build this bridge to implementation.

What Next?

<div style="float:right; border:3px solid black; padding:20px; font-size:48px; font-weight:bold;">8</div>

Building on the Whole-Person Concept

The Army has long held the view that what matters is the whole person. Indeed, the "whole-person" concept has driven the selection and classification of officers for many years. Officer selection boards pore over test scores, school records, and personal recommendations about candidates before making their decisions. But there are far more enlisted applicants than there are prospective officers, and so it has been harder to apply this concept to enlisted volunteers.

This report has been about research to enable Army managers to better evaluate the potential performance of applicants for enlistment, even though there are hundreds of thousands of them each year. The newly developed spatial and psychomotor tests, temperament measures, and interest questionnaires all help to provide a more complete view of an applicant than the ASVAB does by itself.

The whole-person concept applies to the assessment of a soldier's performance as well as to the assessment of an applicant's potential. Here too, Army research has advanced our understanding of the different ways in which that contribution can take form -- for example, through bringing about better "Can-Do" or "Will-Do" aspects of enlisted performance.

Selection as Part of Overall Personnel Management

Selection is only one component of an overall personnel management system that also includes classification, training, and supervision and leadership. These strategies combine to determine the performance and readiness level of the Army enlisted force.

Classification decisions that are based more on the best possible person-job match than on other factors (e.g., solely on the need to fill training seats) will help assure that soldiers entering training have the highest potential for success. But training success does not lie exclusively with the new recruit. The Army must provide well-qualified instructors who have the resources, training techniques, and time to effectively convey to these novice soldiers the skills and motivation they need to prepare them for Army service.

Once on the job, a soldier's learning days are not over. NCO leaders must continue the task of developing job skills and doing what they can to assure that their soldiers adapt to and thrive in the Army environment. At every step of the way, Army leadership decisions on assigning, promoting, and retaining individual soldiers affect the overall performance of the force.

Attending to any one of these components without regard for the contributions of the others would indeed be short-sighted, as well as limiting the potential effectiveness of tomorrow's Army.

The Continuing Need to Learn More

In addition to providing many valuable answers, Project A/Career Force has uncovered important new questions and opportunities. Follow-on work is under way at ARI on more powerful tools for matching persons with jobs. On the performance side, tests of readiness for higher level NCO positions are being explored for possible use in developing enlisted leadership.

The implementation process itself would benefit from additional study, to help technical advances fold seamlessly into a complex Army that works quite well already. The follow-ons to Project A/Career Force will build on its solid base, and they will change the face of selection and classification.

What the Future Holds

After a turbulent half decade of drastic downsizing, the Army's personnel strength is leveling off. But the vision for the future is anything but static. Army leaders have been working hard to prepare for the future and to shape it. The outcome of these efforts is a vision of Force XXI -- the Army of the 21st century.

As the Army adapts to its smaller size, it is likely to require more generalization of skills. Rather than many highly specialized MOS to accommodate more sophisticated equipment, the greater pull will be for less specialization because fewer people are available to do the work. Boundaries between jobs are going to be less well-defined and our forces will be more tailored to specific missions.

Technology will make even more information available at all levels. In the Army, systems such as the Automated Field Artillery Target Delivery System (AFATDS) and the Intervehicular Information System

---

Army Vision--Force XXI

Our 21st Century Army . . . is forming right now. In 2010, our Army's mission will be the same as it has been since 1775 -- to fight and win our Nation's wars, to serve America, to gain decisive success at whatever we are asked to do. We must fulfill that mission today, tomorrow, and the day after tomorrow, in 1994, in 2010, and in every year between.

External events will certainly change our Army over the next 16 years, and we cannot foresee all of them. The Army of 2010 will be different from the Army of 1994. To shape the Army of 2010 we must act now and influence the course our Army pursues. The real challenge is to think beyond the inevitable, and the probable, to take hold of our own destiny.

Army Science and Technology, Master Plan[13]

(IVIS) will deliver vast amounts of information to front-line fighters.[47] This is a far cry from the old approach of "just obey orders and don't think too much."

The nature of war will become even less predictable. Players in the world game proliferate (including non-governmental organizations and ethnic/nationalist groups), and instability can come from any direction. Shifting national alliances and rapid introduction of new technologies will further complicate the Army's ability to anticipate the challenges it will face. Add to this the expectation that the military will carry out diverse operations other than war -- humanitarian aid, peacekeeping -- and it is clear that the Army's future will cover a lot of territory.

Finally, the trend toward conducting operations in more complex settings -- joint, multinational, and United Nations -- is

likely to continue. Army personnel can expect to work more and more with people outside the framework of U.S. Army structure, discipline, and training.

These are among the most significant characteristics of the future anticipated by Army leaders. A common theme is the requirement to adapt -- to a smaller Army, to diverse missions, to increased amounts of information and the expectations that go along with it. More detail on the question of what the future is likely to hold for the Army is provided in TRADOC Pamphlet 525-5.[50]

Implications for Selection and Classification

The vision of Force XXI has implications for all aspects of the Army's plans and operations. In the end, however, the Army must turn to its new accessions to assure its continued well-being and mission success in the future. What types of individuals will be needed? How different are they from those the Army now tries to recruit?

Some speculative answers to these questions were provided to Army leaders at the Army 2010 Conference: Future Soldiers and the Quality Imperative, held in the summer of 1995. These answers were the judgments of ARI scientists with extensive experience in Army selection and classification.[38] Working from a comprehensive list of job-relevant human attributes, these scientists made judgments about the importance of the attributes in performing enlisted jobs in the future that is envisioned by Army leaders.

The attributes that were judged to be the most important for successful performance in the Army of the 21st Century were headed by cognitive ability, integrity, motivation to

---

> **A Vision of the Future**
>
> **Fewer MOS**
> **Information Explosion**
> **Diverse Missions**
> **Quick-Response Work Units**
> **Interagency, Multinational Operations**

---

achieve, cooperativeness, conscientiousness, and adaptability. Also, the ARI judges were asked to identify the attributes most likely to *increase* in importance in the future. In this ranking, cognitive ability dropped only to second place, with the cognitive dimensions of spatial ability, perceptual speed/accuracy, and comprehension monitoring also ranking highly. These judgments suggest that the Army will have to recruit an even higher percentage of bright youth than it does today.

Reflecting the vision of Force XXI in other ways, the ARI judges ranked adaptability as the attribute gaining most in importance, with openness, cultural adaptability, tolerance of ambiguity, ethnic/racial tolerance, and emotional stability also ranking highly. Future research in selection and classification must continue to identify the human traits that will be essential for success in the changing battle environment.

A Never-Ending Process

It should be clear by now that improving the selection and placement of enlisted soldiers is neither a simple nor a finite task. Instead, it is a continuing process of learning, through research and experience, and adapting to changes in Army needs. Those needs, too, continually evolve because of the rapidly changing world around us. A "steady state" is nowhere in sight.

The goal of Army researchers is to work hand-in-hand with Army leaders to monitor the changing environment and to collect and analyze research data in a way that will help respond to the Army's needs. The costs of research seem high, but the investment has a deep and enduring return. Illustrating this point is the return on research investment that ASVAB makes, hundreds of times a day. Working together, Army leaders and the researchers who support them will maintain and improve a proud, effective organization of quality soldiers.

We are making the Army of tomorrow a reality of today. We are meeting the needs of the 21st century by leveraging technology and designing a force that can better accommodate the rapidly changing world situation. . . . In the midst of all this change, however, one constant remains. That constant is the need to fill the ranks of America's Army with quality soldiers. Only intelligent, physically fit, highly motivated, disciplined, and well-trained soldiers can leverage technology to its full potential. Soldiers will be the critical enabler of Force XXI, for it is through quality soldiers that we will realize the full power of technology.

GEN Gordon R. Sullivan
Chief of Staff, United States Army

**The Army
Is
Equipment -**
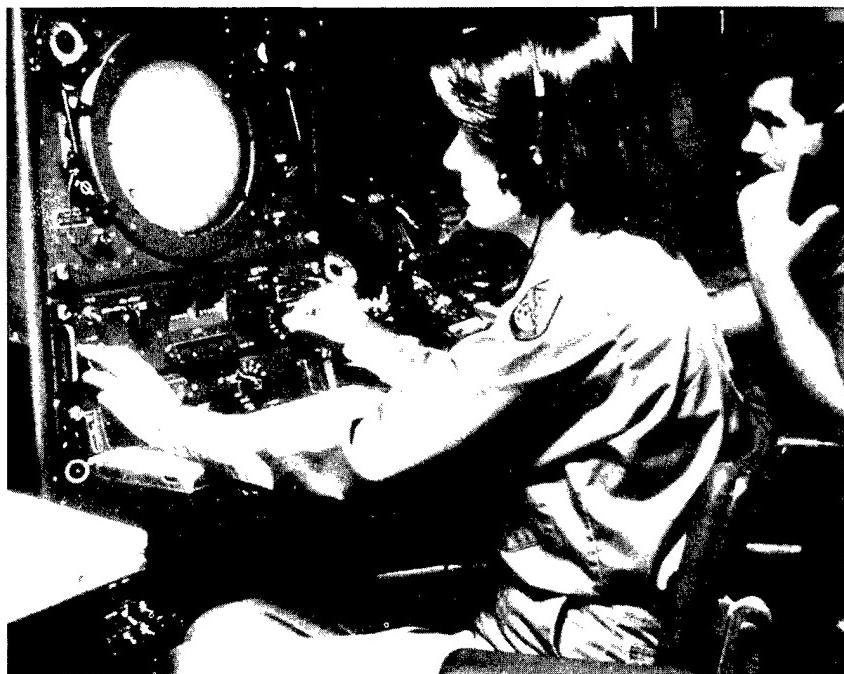
But Above All,

**The Army Is
Human Power**

REFERENCES

1.   Alderton, D. L., Wolfe, J. H., & Larson, G. E. (1996). The Enhanced Computer Administered Test (ECAT) Battery. *Military Psychology, 9* (1).

2.   Alexander, A.. (1980). *Policy research on human issues.* Washington, DC: Army Science Board, Human Issues Group No. 2. Cited in Rumsey and Walker, *The future: A research agenda.*

3.   Block, D. J., Polich, J. M., & Horne, D. K. (1989). *Soldier Performance Research Project.* U.S. Training and Doctrine Command.

4.   Busciglio, H. H. (1990). *The incremental validity of spatial and perceptual-psychomotor tests relative to the Armed Services. Vocational Aptitude Battery* (Technical Report 883). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

5.   Busciglio, H. H., Silva, J., & Walker, C. B. (1990). *The potential of new Army tests to improve job performance.* Paper presented at the Army Science Conference, Durham, NC.

6.   Campbell, J. P. (1987). *Improving the selection, classification, and utilization of Army enlisted personnel: Annual Report, 1986 fiscal year* (Technical Report 813101). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

7.   Campbell, J. P. (1990). An overview of the Army selection and classification project (Project A). *Personnel Psychology, 43* (Special Issue).

8.   Campbell, J. P., McCloy, R. A., Oppler, S. H., & Sager, C. E. (1993). A theory of performance. In N. Schmitt & W. C. Borman (Eds.) *Personnel selection in organizations.* San Francisco: Jossey-Bass.

9.   Campbell, J. P., Russell, T. L., & Knapp, D. J. (1994). *Building an agenda for Joint-Service classification research* (AL/HR-TP-1994-0003). Brooks Air Force Base, TX: Armstrong Laboratory, Air Force Materiel Command.

10.   Campbell, J. P., & Zook, L. M. (Eds.) (1991). *Improving the selection, classification, and utilization of Army enlisted personnel: Final Report on Project A* (Research Report 1597). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

11.   Campbell, J. P., & Zook, L. M. (Eds.) (1996). *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Final Report* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

12.   DeMatteo, J. S., White, L. A., Teplitzky, M. L., & Sachs, S. A. (1991). *Relationship between temperament constructs and selection for Special Forces training,* Paper presented at the Annual Conference of the Military Testing Association, San Antonio.

13.   Department of the Army (1995). *Army Science and Technology Master Plan,* Volume I. Washington, DC: Office of the Assistant Secretary of Army (Research, Development, and Acquisition).

14.   Eaton, N. K., Wing, H., & Mitchell, K. J. (1985). Alternate methods of estimating the dollar value of performance. *Personnel Psychology, 38* (1).

15.   Egbert, R. L., Meeland, T., Cline, V. B., Sprinkler, E. W., & Brown, C. (1958). *FIGHTER I: A study of effective and ineffective combat performers* (Special Report 13). Presidio of Monterey, CA: U.S. Army Leadership Human Research Unit (Human Resources Research Office).

16.   Eitelberg, M. J. (1988). *Manpower for military occupations.* Washington, DC: Office of the Assistant Secretary of Defense (Force Management and Personnel).

17.   Flyer, E. S. (1959). *Factors relating to discharge for unsuitability among 1956 airmen accessions to the Air Force* (WADC-TN-59-201). Lackland Air Force Base, TX: Personnel Research Laboratory.

18.   Grafton, F. C., & Horne, D. K. (1985). *An investigation of alternatives for setting second-to-third tour reenlistment standards* (Technical Report 690). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

19.   Harris, D. A., McCloy, R. A., Dempsey, J. R., DiFazio, A. S., & Hogan, P. F. (1993). 93-35). Alexandria, VA: Human Resources Research Organization.

20.   Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response-distortion on those validities. *Journal of Applied Psychology, 75,* (5), 581-595.

21.   Johnson, C. D., & Zeidner, J. (1990). *Classification utility: Measuring and improving benefits in matching personnel to jobs* (IDA Paper P-2240). Alexandria, VA: Institute for Defense Analyses.

22.   Keene, J. D. (1994). Intelligence and morale in the army of a democracy: The genesis of military psychology during the first World War. *Military Psychology, 6,* (4).

23.   Knapp, D. J., Carter, G. W., McCloy, R. A., & DiFazio, A. S. (1996). The role of job satisfaction in performance, attrition, and reenlistment. In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1993 fiscal year* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

24.   Knapp, D. J., Russell, T. L., & Campbell, J. P. (1993). *Building a Joint-Service classification research roadmap: Job analysis methodologies* (IR-PRD-93-15). Alexandria, VA: Human Resources Research Organization.

25.   Laurence, J. H. (1993). Education standards and military selection. In T. Trent & J. H. Laurence (Eds.), *Adaptability screening for the Armed Forces.* Washington, DC: Office of the Assistant Secretary of Defense (Force Management and Personnel).

26.   Matyuf, M. M., & White, L. A. (1991). *Individual differences in at-fault accident behavior.* Paper presented at the Annual Conference of the Military Testing Association, San Antonio.

27.   McCloy, R. A. (1996). The sequential prediction of individual job performance: The "roll-up." In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Final Report* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

28.   McCloy, R. A., & DiFazio, A. S. (1996). Prediction of first-term military attrition using pre-enlistment predictors. In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1993 fiscal year* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

29.   McCloy, R. A., Ramsberger, P. F., Harris, J. H., Campbell, C. H., & Laurence, J. H. (1992). VA: Human Resources Research Organization.

30.   McHenry, J. J., Hough, L. M., Toquam, J. L., Hanson, M. A., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43* (Special Issue).

31.   Oppler, S. H., Campbell, J. P., & Peterson, N. G. (1996). Maximizing selection validity for predicting first-tour performance. In J. P. Campbell & L. M. Zook, *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Final Report* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

32.   Peterson, N. G., Oppler, S. H., Sager, C. E., & Rosse, R. L. (1994). *Analysis of the Enhanced Computer Administered Test Battery: An evaluation of potential revisions and additions to the Armed Services Vocational Aptitude Battery.* Washington, DC: American Institutes for Research.

33.   Peterson, N. G., Sager, C.E., Oppler, S. H., & Rosse, R. L. (1996). Identification of optimal predictor batteries using a subset of the Project A/Career Force Experimental Predictor Battery. In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1993 fiscal year* (Research Note 96-45). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

34.   Project A: The U. S. Army Selection and Classification Project (1990). *Personnel Psychology, 43* (Special Issue).

35.   Ramsberger, P. F., Laurence, J. H., McCloy, R. A., & DiFazio, A. S. (1994). *Augmented selection criteria for enlisted personnel* (FR-PRD-94-07). Alexandria, VA: Human Resources Research Organization.

36.   Rudnik, R. A., & Greenston, P. M. (1995, 1996). *Development of an Army prototype PC-based Enlisted Personnel Allocating System (EPAS)* (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

37.   Rumsey, M. G. (1994, October). *Project A and Building the Career Force: A summary.*

38.   Rumsey, M. G. (1995). *The best they can be: Tomorrow's soldiers.* Paper presented at The Army 2010 Conference, Cantigny.

39. Rumsey, M. G., & Simutis, Z. M. (1994). *Linking enlisted soldier selection to job performance.* Paper presented at the annual conference of the Military Testing Association.

40. Rumsey, M. G., & Walker, C. B. (1994). The future: A research agenda. In M. G. Rumsey, C. B. Walker, & J. H. Harris (Eds.), *Personnel selection and classification.* Hillsdale, NJ: Lawrence Erlbaum Associates.

41. Russell, T. L., Knapp, D. J., & Campbell, J. P. (1992). *Building a Joint-Service classification research roadmap: Defining research objectives* (IR-PRD-92-10). Alexandria, VA: Human Resources Research Organization.

42. Sadacca, R., White, L. A., Campbell, J. P., DiFazio, A. S., & Schultz, S. R. (1989). *Assessing the utility of MOS performance levels in Army enlisted occupations* (Technical Report 839). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

43. Scholarios, D. M., Johnson, C. D., & Zeidner, J. (1994). Selecting predictors for maximizing the classification efficiency of a battery. *Journal of Applied Psychology, 79* (3), 412-424.

44. Schwartz, A. C., & Mael, F. A. (1992). *An introduction to the Army personnel system* (Special Report S-19). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

45. Shields, J. L., & Baker, J. D. (1981). *The Army's personnel problem and the Golden Spike solution.* Paper presented at the National Security Industrial Association conference, San Diego.

46. Shields, J. L., & Hanser, L. M. (1990). Designing, planning, and selling Project A. *Personnel Psychology, 43* (Special Issue).

47. Smith, D. A., Sticha, P. J., & Morrison, J. E. (1995). *Soldier as adaptive problem solver and decision maker.* Paper presented at the Army 2010 Conference: Future Soldiers and the Quality Imperative.

48. Trent, T., & Laurence, J. H. (Eds.) (1993). *Adaptability screening for the Armed Forces.* Washington, DC: Office of the Assistant Secretary of Defense (Force Management and Personnel).

49. U.S. Army Recruiting Command (1995, June). *Recruiting America's Army in the twenty-first century.* The Army 2010 Conference, Cantigny.

50. U.S. Training and Doctrine Command (1994). *Force XXI Operations* (Pamphlet 525-5). Fort Monroe, VA: Author.

51. Walker, C. B. (1989). *The progress and results of predictor testing in TRADOC's Skills Selection and Sustainment (S3) Program* (WP-RS-89-23). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

52. White, L. A., Mael, F. A., & Sachs, S. A. (1991). *Selection and candidates for Ranger training.* Unpublished manuscript. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

53. White, L. A., Nord, R. D., Mael, F. A., & Young, M. C. (1993). The Assessment of Background and Life Experiences (ABLE). In T. Trent & J. H. Laurence (Eds)., *Adaptability screening for the Armed Forces.* Washington, DC: Office of the Assistant Secretary of Defense (Force Management and Personnel).

54. White, L. A., Rumsey, M. G., & Matyuf, M. M. (1994). *Relationships between soldiers' performance during peacetime and in combat.* Paper presented at the American Psychological Association convention, Los Angeles.

55. Wise, L. L., Arabian, J. M., Chia, W. J., & Szenas, P. L. (Eds.) (1989). *Army Synthetic Validity Project: Report of Phase I Results.* Washington, DC: American Institutes for Research.

56. Wise, L. L., Peterson, N. G., Hoffman, R. G., Campbell, J. P., & Arabian, J. M. (Eds.) (1991). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

57. Wiskoff, M. F., Parker, J. P., Zimmerman, R. A., & Sherman, F. (1989). *Predicting school and job performance of Marine Security Guards* (PERSEREC Technical Report 90-003). Monterey, CA: Defense Personnel Security Research and Education Center.

58. Zeidner, J., & Drucker, A. J. (1988). *Behavioral science in the Army: A corporate history of the Army Research Institute.* Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

59. Zeidner, J., & Johnson, C. D. (1989). *The economic benefits of predicting job performance*

60. Zeidner, J., & Johnson, C. D. (1989). *The utility of selection for military and civilian jobs* (IDA Paper P-2239). Alexandria, VA: Institute for Defense Analyses.

Additional Sources

Campbell, C. H., Campbell, R. C., Rumsey, M. G., & Edwards, D. C. (1986). *Development and field test of task-based MOS-specific criterion measures* (Technical Report 717). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Campbell, C. H., Ford, P., Rumsey, M. G., Pulakos, E. D., Borman, W. C., Felker, D. B., De Vera, M. V., & Reigelhaupt, B. J. (1990). Development of multiple job performance measures in a representative sample of jobs. *Personnel Psychology, 43* (Special Issue).

Campbell, J. P., McHenry, J. J., & Wise, L. L. (1990). Modeling job performance in a population of jobs. *Personnel Psychology, 43* (Special Issue).

Campbell, J. P., & Russell, T. L. (Eds.) (1994). *Building a Joint-Service classification research roadmap: Methodological issues in selection and classification issues* (AL/HR-TP-1994-0013). Brooks Air Force Base, TX: Armstrong Laboratory, Air Force Materiel Command.

Graham, S. E. (1989). *Assessing the impact of mental category on simulated tank gunnery performance* (Research Report 1515). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Graham, S. E. (1989). *The effects of soldier quality on armor field and Simnet test performance.* Fort Knox, KY: U.S. Army Research Institute Fort Knox Field Unit.

Green, B. F., Jr., & Wing, H. (Eds.) (1988). *Analysis of job performance data: Report of a workshop,* Committee on the Performance of Military Personnel. Washington, DC: National Academy Press.

Hough, L. M. (1988). *Literature review: Utility of temperament, biodata, and interest assessment for predicting job performance* (Research Note 88-02). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Knapp, D. J., & Campbell, J. P. (1993). *Building a Joint-Service classification research roadmap: Criterion-related issues* (AL/HR-TP-1993-0028). Brooks Air Force Base, TX: Armstrong Laboratory, Air Force Materiel Command.

McHenry, J. J., & Rose, S. R. (1988). *Literature review and potential usefulness of psychomotor ability tests for selection and classification* (Research Note 88-13). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Oppler, S. H., Campbell, J. P., & Peterson, N. G. (1996). Maximizing selection validity for predicting first-tour performance. In J. P. Campbell & L. M. Zook, *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Final Report* (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Peterson, N. (Ed.) (1986). *Development and field test of the Trial Battery for Project A* (Technical Report 739). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Peterson, N. G., Hough, L. M., Dunnette, M. D., Rosse, R. L., Houston, J. S., & Toquam, J. L. (1990). Project A: Specification of the predictor domain and development of new selection/classification tests. *Personnel Psychology, 43* (Special Issue).

Peterson, N., Russell, T., Hallam, G., Hough, L., Owens-Kurtz, C., Gialluca, K., & Kerwin, K. (1990). Analysis of Experimental Predictor Battery: LV sample. In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel - Annual Report, 1990 fiscal year* (Technical Report 952). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Pulakos, E. D., & Borman, W. C. (1987). *Developing the basic criterion scores for Army-wide and MOS-specific ratings* (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Rumsey, M. G., Walker, C. B., & Harris, J. H. (Eds.) (1994). *Personnel selection and classification.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Russell, T. L., Crafts, J. L., Tagliareni, F. A., McCloy, R. A., & Barkley, P. (1994). *Job analysis of Special Forces jobs* (Technical Report). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Russell, T. L., Reynolds, D. H., & Campbell, J. P. (1994). *Building a Joint-Service classification research roadmap: Individual differences measurement* (AL/HR-TP-1994-0009). Brooks Air Force Base, TX: Armstrong Laboratory, Air Force Materiel Command.

Sadacca, R., Campbell, J. P., DiFazio, A. S., Schultz, S. R., & White, L. A. (1990). Scaling performance utility to enhance selection/classification decisions. *Personnel Psychology, 43*

Sticha, J. P. (1994). New methods for estimating gains due to classification and new procedures for making differential job assignments. In J. P. Campbell (Ed.) & T. L. Russell, *and classification* (AL/HR-TP-1994-0013). Brooks Air Force Base, TX: Armstrong Laboratory, Air Force Materiel Command.

Toquam, J. L., Corpe, V. A., & Dunnette, M. D. (1989). *Literature review: Cognitive abilities -- theory, history, and validity* (Research Note 91-28). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Walker, C. B. (1986). *The Army's Military Applicant Profile: Its background and progress* Defense). In symposium at the Annual Conference of the Military Testing Association, Munich.

White, L. A. (1994). Development of composite scores for Assessment of the Background and Life Experiences (ABLE) instrument. In J. P. Campbell & L. M. Zook (Eds.), *Building and retaining the Career Force: New procedures for accessing and assigning Army enlisted personnel: Annual Report, 1991 fiscal year* (Research Note 94-10). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wise, L. L., McHenry, J. J., & Campbell, J. P. (1990). Identifying optimal predictor composites and testing for generalizability across jobs and performance factors. *Personnel Psychology, 43* (Special Issue).

Wise, L. L., Peterson, N. G., Hoffman, R. G.., Campbell, J. P., & Arabian, J. M. (1991). *Army Synthetic Validity Project: Report of Phase III Results. Vol. II: Research Instruments.* (Research Product 91-08). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Wolfe, J. H. (in press). Incremental validity of Enhanced Computer Administered Test factors. *Military Psychology.*

Young, W. Y., Houston, J. S., Harris, J. H., Hoffman, R. G., & Wise, L. L. (1990). Large-scale predictor validation in Project A: Data collection procedures and data base preparation. *Personnel Psychology, 43* (Special Issue).

GLOSSARY

Ability

A human trait that makes a person more or less capable of performing various functions. An ability is relatively stable in that it cannot be improved quickly through training. Examples: general cognitive ability, spatial ability, psychomotor ability.

AFQT

(See below Armed Forces Qualification Test)

Aptitude

One's potential to learn particular types of behavior or content. Equated by many with "trainability." Contrasted with other types of human attributes that are dispositions, such as temperament and interests, rather than capacities or potential.

Aptitude Area Composites

The scores used by the Army for classifying enlistees into MOS (jobs). They are computed by adding scores on specific subsets of ASVAB subtests. Examples are Clerical, Mechanical, Maintenance, Electronics, and Combat.

Archival Records

Army personnel records and databases, such as Enlisted Efficiency Reports and the Enlisted Master File, that routinely store indicators of individuals' demographics and success on the job. Some of the indicators to be found in these records are promotions, disciplinary actions, and awards.

Armed Forces Qualification Test (AFQT)

A score used by all of the Services for selecting volunteers for enlistment. It is computed by adding the scores of these ASVAB subtests: Word Knowledge, Paragraph Comprehension, Arithmetic Reasoning, and Math Knowledge. Applicants are grouped into categories (I, II, IIIA, IIIB, IV, or V) based on their AFQT score.

Armed Services Vocational Aptitude Battery (ASVAB)

The battery of 10 subtests given to all applicants for enlisted military service. A composite of four of those subtests makes up the Armed Forces Qualification Test (AFQT).

Attrition

Failure to complete the first enlistment successfully. Of concern to the Services because of the cost of recruiting, training, and replacing individuals who do not go on to return that investment by performing successfully.

Batch A; Batch Z

In Project A/Career Force, the nine Batch A MOS were given all of the performance measures (see Figure 3.2) in their first and second enlistment. Performance in Batch Z MOS was evaluated only as far as the first enlistment, and not on MOS-specific hands-on nor job knowledge measures.

Biodata

A self-report (i.e., questionnaire) predictor of job performance that asks about one's life experiences (e.g., involvement in extracurricular activities, work experience).

"Can-Do" Performance

The skilled performance that one is capable of, as contrasted with the performance that one typically exhibits (i.e., *"Will-Do"*). Both "Can-Do" and "Will-Do" performance are necessary for full success on the job.

CAST

The Computerized Adaptive Screening Test, a short test of Word Knowledge and Arithmetic Reasoning. Administered at Recruiting Stations to identify applicants who are likely to qualify on the full AFQT (see Figure 1.1).

Category I...V

Intervals on the scale of scores on the AFQT (see Table 1.2). These levels of AFQT are used as an indicator of applicants' level of trainability.

Classification

Person-job matching, or putting the right person in the right job. The process and technology of assigning persons to the jobs for which they have the prerequisite characteristics of intellect and/or temperament. For assigning civilian volunteers to Army jobs, classification is based on the individuals' scores on the Aptitude Area Composites of ASVAB. This meaning, also referred to as personnel classification, contrasts with *occupational classification*, which deals with the grouping of jobs, as into Army Career Management Fields or Branches.

Compensatory Screen

A system for selecting new enlistees that permits strengths in one area to offset (compensate for) weaknesses in another. For example, higher AFQT scores are allowed to compensate for applicants' not having a high school diploma.

Composite

A single score or measure that is formed by combining two or more other scores. In Project A, an individual's score on MOS-Specific Skill was a composite of the person's scores on hands-on testing and pencil/paper knowledge testing. In turn, each of those two components was a composite of scores on numerous tasks or test items.

Correlation (Correlate, Correlation Coefficient, r)

The degree to which persons' scores on one measure change (or covary) with their scores on another measure. For example, measures of persons' heights and weights are directly correlated, but there is almost no correlation between persons' intelligence and their degree of social outgoingness. The relationship between two sets of measures is summarized numerically in a correlation coefficient (r) that can range from +1.0 (a positive one-to-one relationship), through 0 (no relationship), to -1.0 (an inverse one-to-one relationship). In evaluating new tests for selection and classification, we look at how strongly they correlate with measures of success on the job. For that purpose, a correlation of .3 or higher (or -.3 or lower) has usually been considered good; but that evaluation depends on a host of considerations.

Criterion

An indicator of success on the job, usually expressed as a number. Project A constructed several composite criterion measures, including the five performance factors (e.g., Core Technical Proficiency), combinations of factors (e.g., "Can-Do" and "Will-Do"), and attrition.

Delayed Entry Program (DEP)

Between the time civilians sign an enlistment contract and come on active duty, which may be as long as a year, they are in this program. The DEP helps the Army to recruit by permitting applicants to coordinate entry on duty with other important life events (e.g., anticipated graduation from high school).

ECAT

The Enhanced Computer Administered Testing project of the Department of Defense in which new computerized tests of the services were evaluated on samples of Army, Navy, and Air Force trainees.

EST

The Enlistment Screening Test, a pencil-and-paper test of Word Knowledge and Arithmetic Reasoning. Administered by Recruiters to determine an applicant's probability of passing the full AFQT (see Figure 1.1).

Fairness in tests

Statistical evidence showing that the performance of minorities or females is not underpredicted by test scores relative to the performance of majority group members or males. If such underprediction were evident, the use of such a test would unfairly discriminate against these groups.

Incremental Validity

A gain in *validity* by adding a new test to another, baseline test. In our case, found when a new test adds to the known strong *predictive* baseline of *ASVAB*.

Maximal Performance

How well one can do a job, equivalent to can-do performance. Measured by one-shot tests (e.g., hands-on performance, written job knowledge).

MOS

Military Occupational Specialty, or an enlisted soldier's Army job. Examples: 11B, Infantryman; 44B, Machinist; 73C, Finance Specialist; and 82C, Field Artillery Surveyor.

Longitudinal Validation (LV)

The phase of Project A/Career Force that evaluated two sets of predictors -- the new project predictor tests and ASVAB -- by testing new recruits, measuring their performance in Army jobs during their first and second enlistments, and then seeing how well the predictors distinguished the examinees' future levels of Army performance (see Figure 2.2).

NCO

Non-commissioned officer. Ranking from Sergeant up through Sergeant Major, NCOs are the enlisted leadership of the Army.

Noncom

See NCO.

Performance Factor

In the Project A research, five critical aspects of soldier performance were identified: General Soldiering Proficiency, Core Technical (MOS-specific) Proficiency, Effort and Leadership, Personal Discipline, and Physical Fitness/Military Bearing. Scores on each of these factors were derived by adding together the appropriate subscores from a variety of individual measures (e.g., peer and supervisor ratings, hands-on tests, written exams).

Person-Job Matching

Determining the job for which an individual is best suited. See Classification.

Prediction

Estimating persons' future success on the job from evidence that is available before they are given the job. In the Army, ASVAB test scores and high school diploma status are the primary predictors of future success.

Predictor

A test or other indicator used as a basis for selecting among applicants for enlistment and for assigning them specific Army jobs. Examples: ASVAB, high school diploma status.

Psychomotor Ability

Ability to perform voluntary movements well (e.g., with speed, accuracy, coordination, and timing).

Quality

Refers to individuals' potential to perform well in the military. Traditionally defined by high school diploma status and AFQT Category. Research in selection and classification is aimed at expanding the concept of quality toward including the whole person.

Reliability (Reliable)

Consistency in measuring a stable human attribute, as from one occasion to another, from one form of a test to another, or from one rater of performance to another. Unless a measure is sufficiently reliable, it cannot be valid. As opposed to interviews and expert judgments, good tests are the most reliable measures of human aptitudes, temperament, and performance.

Selection

Determining basic eligibility for enlistment or advancement. Screening is an equivalent term, these contrasting with the other major use of tests in personnel processing: classification.

Spatial Ability

The ability to use information on the relations of things in space such as direction, form, rotation, position in a layout, path, orientation, etc.

SPRP

Soldier Performance Research Project. An evaluation of ASVAB in terms of levels of soldiers' performance on specific tasks and jobs as a function of their scores on AFQT. Found strong positive relations between AFQT and job performance (see Figure 6.4).

SQT

Skill Qualification Test, a test of MOS-specific skills given periodically to soldiers in most MOS to determine their continuing proficiency in their specialty. Discontinued in 1991.

Typical Performance

The level of performance usually exhibited on the job. Not measurable using a one-shot test, it is measured best by the ratings of persons who have had the opportunity to see each ratee perform on a day-to-day basis

Utility

The value to the Army of a test or procedure. Sometimes quantifiable (e.g., in dollars, first round hits, number of personnel needed, change in rate of attrition). The net in a calculation of costs and benefits.

Validity (Validate, Validation)

Suitability of a test for the purposes it is intended to serve. There are three basic strategies for demonstrating validity. Content validity is shown by evidence that the content of the test matches the content which it is supposed to assess. Criterion-related validity of a predictor is shown when individuals' scores on it correlate highly enough with their scores on a relevant criterion. What psychological trait or attribute the test measures is the issue in construct validity. Researchers try to settle that issue with evidence indicating that scores on the test are dependably related with scores on other types of relevant measures, results of experiments, and other evidence.

"Will-Do" Performance

The motivational component of job performance, including effort, discipline, fitness, and military bearing. Contrasts with *"Can-Do"* performance. Both "Can-Do and "Will-Do" performance are required for full success on the job.

APPENDIX A: Phases in the Army Life
of a Career Soldier

Testing/Selection/Classification

**Army
Recruiting Station**

DoD/Army
MEPS: Complete Selection and Classification

Pass
Prescreen
(CAST/EST) — Yes → Pass
ASVAB
Testing — Yes → Pass
Physical,
Other — Yes → Guidance
Counselor → Classification/
Contract →

(Continued)

Army Enlistment   *Training*

→ Enter Delayed
Entry Program
(DEP) → Accession → Reception
Processing → Basic
Training → Limited
Classification → Advanced
Individual
Training →

Discharge
from DEP
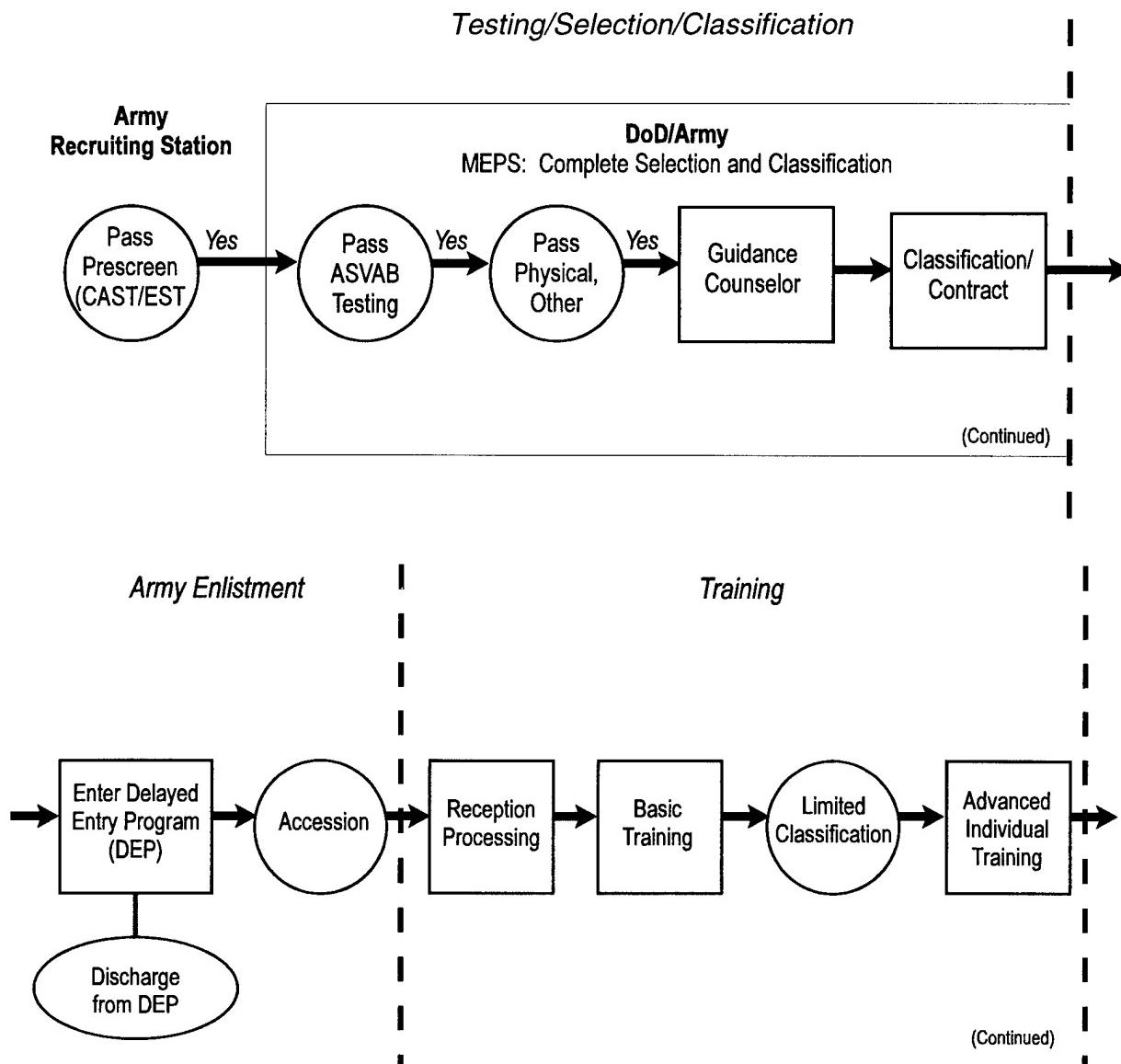
(Continued)

Phases in the Army Life of a Career Soldier (Continued)

First Term/Post-Training

Assignment

Special Programs

- - - - - - - - -

Example: Rangers

Gain Promotion?
E3/E4/E5

Complete
First
Term?

(Continued)

Meet Minimum
Grade Requirement?
(Varies by MOS)

Pass
Physical?

Pass
Water Survival?

Meet Medical
Requirements?

Pass ASVAB
Screen?

Apply for
Ranger
Training

Complete
Ranger
Training

Receive Duty
Assignment

Meet
Time-in-Grade Minimum
Service Requirement?

Meet
Minimum Service
Requirement?

Meet
Promotion Point
Cutoff?

No
Flags?

Recommended
by Commander

Approved by
Board

Approved by
Promotion
Authority?

Promoted to
Sergeant

Phases in the Army Life of a Career Soldier (Continued)

Second Term

Reenlist ?	Gain Promotion? E4/E5	Complete Second Term ?

Apply to Reenlist ?	Bar to Reenlist?	E3 or Above	Eligible for Discharge	Eligible for Chosen Option?	Reenlistment Approved

No — Yes — Yes — Yes
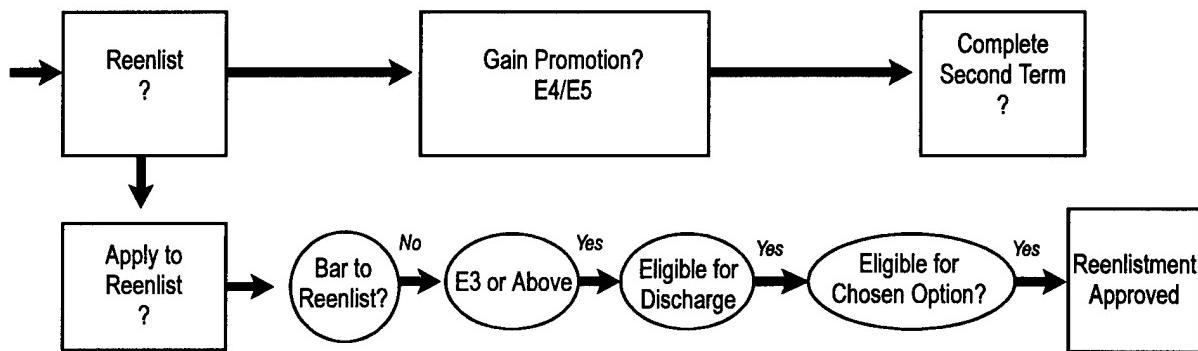
APPENDIX B: A Brief History of Military Selection

Who Makes a Good Soldier?

That question has been asked for a very long time. Formal testing to select and classify personnel may well have had its beginning some 3,000 years ago in Imperial China. In 1115 B.C., a system of competitive examinations, using standardized measures of individual differences, led to appointments in the imperial bureaucracy. Before long, the system expanded to choose individuals for particular military specialties. Spear throwers, for example, were selected on the basis of what would now be called standardized measures of long-distance visual acuity, such as identifying stars in the night sky.[11] Over the centuries, however, military managers have generally coped as best they could to get the soldiers they needed -- militia, conscripts, mercenaries, volunteers -- to fight their battles.

During the American Revolution the colonial forces pretty much accepted anyone who volunteered, as long as he could "walk, talk, see, and hear."[16] As early as 1814, the Army did use examinations to test surgeons and prospective Military Academy trainees.[58] More typically, during the Mexican War of 1847, Army enlistment was limited to "all free white male persons, above the age of 18, and under 35 years, being at least 5 feet 3 inches high, who are effective, able-bodied, sober, free from disease, and who have a competent knowledge of the English language."[16] However, more specialized Army roles began to evolve in the Civil War. Skilled craftsmen were needed for jobs such as artificer, saddler, farrier, cooper, wagoner, and carpenter. These men were in short supply

---

Choosing the "Guardians"

"In no form of work is efficiency so important as in war," Plato observes in his greatest dialogue, *Republic.* Indeed, the conduct of war is "art," he states, where the work of "Guardians," or soldiers of the state, is "the most important of all." It follows, then, that efficient soldiers need to have proper training, because "just taking up a shield or other weapon will not make a man capable of fighting." Guardians, of all workers, require "the greatest amount of skill and practice." But, even more important, Plato writes, they should have "a native aptitude for their calling":

> So, it is our business to define, if we can, the natural gifts that fit men to be guardians of a commonwealth, and to select them accordingly. It will certainly be a formidable task; but we must grapple with it to the best of our power.

Plato recognized the problem of variation in individual endowments, and he may have been the first to suggest that an aptitude test be devised to select persons for soldierly service. That was in the 4th Century B.C. It would take another 2,300 years or more before the first pencil and paper tests were used to screen prospective guardians for any special talents or "natural gifts." Even then, scientists, policymakers, and philosophers of a different order would continue to grapple with the formidable task of finding fit warriors.

Manpower for Military Occupations[16]

and high demand by the Army, and were treated differently from the common soldier.[15]

Starting to Provide a Reasonable Basis for Selection: World War I

The military is unique in needing to make large numbers of personnel decisions fast. When war was declared in April 1917, the Regular Army had about 127,000 officers and enlisted men, with another 80,000 National Guardsmen in federal service along the Mexican border. In the next year and a half, Federal armed forces would grow to 3.6 million men, 72 percent of them drafted and 2.4 million of them in the Army. This wildly varied mix of civilians had to be placed not only in the traditional combat units but also in the vast new array of technical and support units that would make up 60 percent of this new kind of Army.[22]

The effort to approach these decisions in a professional way surfaced in World War I, when the Army faced massive mobilization. Civilian managers had turned to the young field of psychology, a specialty that had come of age since the turn of the century as an expanding America sought scientific management. Still on trial in the academic community, psychologists had been successful in commercial settings. Their experience in helping industry (e.g., General Electric, Westinghouse) to put "the right man in the right job" became the basis for Army tests to choose and structure a vast new military force.[22]

Three types of tests (growing out of Binet's work to test individual differences, or "intelligence") were developed: the written, group-administered Army Alpha for literate recruits, the pictorial Army Beta for illiterates and for men who failed the Alpha, and individual assessment for men who failed the Beta. Scores were made available to

commanding officers for use in assigning individuals to specific duties, selecting noncommissioned officers, and balancing the "intelligence" of military units.

Although the test scores were not accepted by all Army officers, they were important in many manpower decisions. For example, almost 8,000 recruits were tagged for immediate discharge as mentally incompetent, and another 8,000 were assigned to special labor duty.[58] By January 1919, the tests had been given to 1,726,966 men, including 42,000 officers; 83,500 of the enlisted men had also been examined individually.

The Army's psychologists did more than conduct tests; they contributed a rough grading system indicating ability to learn, oral assessments of job knowledge, job specifications, an officer qualification study, data on illiteracy rates, advice on training and morale, and other early studies on "issues that would become staples of 20th century military psychology," such as how recruits adjusted to Army life.[58,22]

An enduring effect of mass testing in World War I was the impetus given to mental ability testing itself. Military interest in tests to classify personnel barely survived the war (both the Army and Navy dropped these testing programs early in 1919), but in civilian life both industrial and educational sources made wide use of mental tests in the decades between the wars.[58]

World War II and Beyond

With the approach of World War II, the Army moved to revive the testing program. Two major influences were the extensive progress that had been made in civilian testing, and the great surge in technical specialization in Army jobs stretching back to the early tanks and planes of World War I.

Well before the Selective Service Act was passed in 1940, the Army had begun a coordinated program to use the tools of psychology in preparing for the massive personnel problems that lay ahead. That prospect became reality in the largest personnel system in history.[58]

Early on, screening was intended mainly to rule out men who might be bad risks, and to sort recruits by their ability to learn quickly; one requirement called only for "the ability to comprehend simple orders given in the English language." The initial new tests were "geared to distinguish those who could read, write, and do sums at a fourth grade level from those who could not"; soon tests were added to select illiterates with enough mental capacity to absorb Army training. As the buildup progressed, specialized tests were added to select soldiers for technical jobs such as mechanic or code clerk,[58,16] and to measure proficiency after training.

These outcomes were primarily the work of research psychologists in the Personnel Research Section [forerunner of the Army Research Institute for the Behavioral and Social Sciences (ARI)], established in the summer of 1940. The PRS staff tried out almost every method and technique that might be used for selection.[58] Much new ground was broken. There were important advances in measuring and analyzing job performance, rating scales were improved, forced-choice and critical-incident methods were tried out, test results were compared with training results and combat records. Much of the research success in the next few decades grew out of the intensive and innovative work during the war years.[1]

Pioneer Testing

Army Alpha	_Army Beta_
Grammar	Mazes
Vocabulary	Block counting
Arithmetic	Number similarities
Analogies	Figure similarities
Common sense	What's missing in this drawing?
Filling in numbers in a sequence	
Unscrambling sentences	

From 1940 to 1945, the Army General Classification Test (AGCT) was used to test some 4,000 persons a day.[58] After the war, when each Service was devising its own aptitude tests, the Army stayed with the AGCT. The Armed Forces Qualification Test (AFQT), introduced in 1950, was modeled after the AGCT but was designed as a screening device that set up "minimum aptitude standards" and established high and low aptitude categories.

In 1974 the Department of Defense directed all the Services to use a single test battery, the Armed Services Vocational Aptitude Battery (ASVAB), which includes the subtests that make up the AFQT. Updated several times, the ASVAB remains the test battery that is universally used for military screening and job assignment.

The last soldier drafted into the U.S. Army reported for training in June 1973, and on July 1 the All-Volunteer Army became official. Between 1940 and the end of the Vietnam War in 1973, more than 14.9 million American youth had been drafted.[46]

The new volunteer system was undertaken with the expectation that difficult times lay ahead with the change from conscription to free choice. By 1975 first-term attrition had reached 27 percent among enlistees who were high school graduates and 51 percent among non-graduates -- both record highs. Although the size of the Army had been drastically reduced from the Vietnam era, these high attrition rates placed a heavy burden on recruiting.[46] The continuing efforts to stabilize and improve the system provided the base for the comprehensive research programs undertaken in the 1980s.

Selection on a Grand Scale

World War I Army Strength -

Year	Strength
1916	108,400
1917	421,500
1918	2,395,700
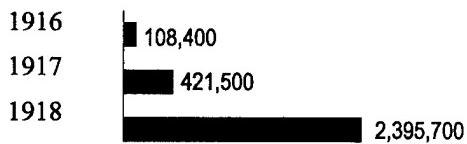
World War II Army Strength -

Year	Strength
1939	190,000
1940	269,000
1945	8,268,000

1   2   3   4   5   6   7   8   9
millions